

JACQUELINE ANDREUCCI LINDSTRON

AS TÉCNICAS HOLÍSTICA E ANALÍTICA NA AVALIAÇÃO DA EXPRESSÃO ESCRITA

Dissertação para a obtenção do Grau de
Mestre. Área de Concentração: Língua
Inglês, do Curso de Pós-Graduação em
Letras, Setor de Ciências Humanas, Letras
e Artes da Universidade Federal do Paraná.

CURITIBA
1997

ZZ-9057

PROFESSOR ORIENTADOR

**Prof. Guido Irineu Engel
Doutor em Sociologia e Professor de
Estatística Aplicada a Estudos de
Linguagem do curso de
Pós- Graduação em Letras da
Universidade Federal do Paraná.**

PROFESSOR CONSULTOR

Prof. Michael Alan Watkins
Professor Assistente da Universidade
Federal do Paraná.

AGRADECIMENTOS

Este trabalho de pesquisa só se tornou possível graças à ajuda de inúmeras pessoas que contribuíram, de uma forma ou de outra, para sua realização. Serei eternamente grata aos alunos que produziram as redações que constituíram o meu instrumento de pesquisa e aos meus colegas que corrigiram estas redações. Em especial, agradeço também à Professora Sílvia Emiko Shimakura, do Departamento de Estatística da UFPR, que me ajudou a organizar os dados e a realizar a análise dos mesmos; e aos Professores Guido Irineu Engel e Michael Alan Watkins por suas inúmeras contribuições no decorrer do trabalho. Não posso aqui deixar de agradecer o apoio e carinho de meu marido, Paulo Roberto, e minha mãe, Inês, que fizeram muito mais do que me incentivar nestes últimos três anos. Sem todas estas pessoas, com certeza, este trabalho de pesquisa não teria se concretizado.

Muito obrigada a todos vocês.

SUMÁRIO

SUMÁRIO.....	i
LISTA DE TABELAS.....	iv
LISTA DE GRÁFICOS.....	v
LISTA DE QUADROS.....	vi
RESUMO.....	vii
ABSTRACT.....	viii
1 INTRODUÇÃO.....	1
1.1 CONTEXTO DO PROBLEMA.....	1
1.2 FORMULAÇÃO DO PROBLEMA E SUA DELIMITAÇÃO.....	1
1.3 JUSTIFICATIVA DO PROBLEMA.....	2
1.4 ESQUEMA DE APRESENTAÇÃO DO TRABALHO.....	2
2 A ESCRITA NO CONTEXTO DO ENSINO DA LÍNGUA ESTRANGEIRA.....	4
2.1 VISÃO GERAL.....	4
2.2 O CONCEITO DE ESCRITA.....	4
2.3 EVOLUÇÃO DO ENSINO DA ESCRITA.....	6
2.3.1 O MÉTODO DA GRAMÁTICA E TRADUÇÃO.....	6
2.3.2 O MÉTODO DIRETO.....	7
2.3.3 O MÉTODO DA LEITURA.....	9
2.3.4 O MÉTODO AUDIOLINGUAL.....	10
2.3.5 A TEORIA COGNITIVA.....	12
2.3.6 A ABORDAGEM COMUNICATIVA.....	15
2.4 CONSIDERAÇÕES FINAIS.....	16
3 A AVALIAÇÃO DA EXPRESSÃO ESCRITA.....	17
3.1 VISÃO GERAL.....	17
3.2 TIPOS DE TESTE.....	17
3.2.1 TESTE "DISCRETE POINT ".....	17

3.2.2 O TESTE INTEGRATIVO.....	18
3.2.3 O TESTE DIRETO <i>VERSUS</i> INDIRETO.....	19
3.2.4 O TESTE COMUNICATIVO.....	19
3.3 CARACTERÍSTICAS ESSENCIAIS DE UM BOM TESTE.....	21
3.3.1 VISÃO GERAL.....	21
3.3.2 PRATICALIDADE.....	21
3.3.3 CONFIABILIDADE.....	21
3.3.4 VALIDADE.....	23
3.4 CONCEITO DE REDAÇÃO COMO INSTRUMENTO DA EXPRESSÃO ESCRITA.....	24
3.5 TIPOS DE REDAÇÃO.....	25
3.6 TÉCNICAS DE AVALIAÇÃO DA REDAÇÃO.....	27
3.6.1 VISÃO GERAL.....	27
3.6.2 A TÉCNICA ANALÍTICA.....	27
3.7.3 A TÉCNICA HOLÍSTICA.....	31
3.7 CONSIDERAÇÕES FINAIS.....	36
4 METODOLOGIA.....	38
4.1 AMOSTRA.....	38
4.2 PROCEDIMENTOS PARA A COLETA DE DADOS.....	38
4.2.1 COM OS ALUNOS.....	38
4.2.2 COM OS PROFESSORES.....	39
4.3 ANÁLISE DOS DADOS.....	41
5 ANÁLISE E INTERPRETAÇÃO DOS DADOS.....	42
5.1 VISÃO GERAL.....	42
5.2 APRESENTAÇÃO DOS DADOS.....	42
5.3 A CORRELAÇÃO.....	45
5.3.1 COEFICIENTE DE CORRELAÇÃO ENTRE OS ESCORES DA TÉCNICA HOLÍSTICA E ANALÍTICA PARA CADA PROFESSOR.....	46

5.3.2 COEFICIENTE DE CORRELAÇÃO ENTRE OS ESCORES DA TÉCNICA ANALÍTICA E HOLÍSTICA PARA TODOS OS PROFESSORES.....	52
5.4 AS MÉDIAS.....	53
5.5 ANÁLISE DE VARIÂNCIA.....	57
5.6 CONSIDERAÇÕES FINAIS.....	63
6 CONCLUSÃO.....	64
ANEXO 1 - AS REDAÇÕES.....	70
ANEXO 2 - INSTRUÇÕES PARA OS PROFESSORES.....	91
ANEXO 3 - ESCALA DE CRITÉRIOS DENOMINADA <i>PMP2</i>	96
ANEXO 4 - ESCALA DE CRITÉRIOS PARA O <i>IELTS</i>	99
ANEXO 5 -. GRÁFICOS DAS NOTAS ATRIBUÍDAS PELA TÉCNICA HOLÍSTICA E ANALÍTICA PARA OS PROFESSORES 2, 3, 4, 5, 6, ,7, 8, 9, E 10.....	102
ANEXO 6 - TABELA DAS MÉDIAS ATRIBUÍDAS PELOS PROFESSORES SEGUNDO A TÉCNICA HOLÍSTICA.....	108
ANEXO 7 - TABELA DAS MÉDIAS ATRIBUÍDOS PELOS PROFESSORES SEGUNDO A TÉCNICA ANALÍTICA.....	110
ANEXO 8 - MODELO DA ANÁLISE DE VARIÂNCIA	112
ANEXO 9 - SOMA DOS QUADRADOS.....	114
ANEXO 10 - CÁLCULO DO QUADRADO MÉDIO.....	116
ANEXO 11 - TESTE F DE SNEDECOR PARA OS TERMOS DO MODELO.....	118
REFERÊNCIAS BIBLIOGRÁFICAS.....	120

LISTA DE TABELAS

1 - Notas atribuídas pelos 10 professores às 20 redações pela Técnica Holística.....	43
2- Notas atribuídas pelos 10 professores às 20 redações pela Técnica Analítica.....	43
3- Notas segundo a Técnica Analítica transformadas para valores percentuais.....	44
4 - Notas e postos para as 20 redações do Professor 1.....	46
5 - Cálculo do "d'e "d ² " para as 20 redações do Professor 1.....	48
6 - Resultados dos Coeficientes de correlação por Postos de Spearman para os dez professores.....	49
7 - Resultado do Coeficiente de Correlação por Postos de Spearman para todos os professores juntos.....	53
8 - Média da nota geral por cada professor.....	55
9 - Médias gerais por Redação.....	56
10 - Análise de Variância para o Modelo.....	59

LISTA DE GRÁFICOS

1 - Notas pelas técnicas holística e analítica para o Professor 1.....	51
2 - Notas pelas técnicas holística e analítica para todos os professores.....	52
3 - Média da Nota Geral pelas Técnicas Holística e Analítica.....	53
4 - Média da Nota Geral dada por cada professor.....	54
5 - Nota Geral atribuída a cada redação.....	56
6 - Interação entre Professor e Técnica.....	60
7- Interação entre Redação e Técnica.....	61
8 - Notas das redações (em ordem de 1 a 20), atribuídas por todos os professores.....	62

LISTA DE QUADROS

1 - Critérios de avaliação do <i>First Certificate in English</i>	29
2 - Critérios de avaliação do <i>Certificate of Proficiency in English</i>	29
3 - Critérios para a correção analítica, de Avosani.....	30
4 - Escala holística utilizada pela <i>Royal Society of Arts Examination in the Communicative Use of English as a Foreign Language</i>	32
5 - Critérios de avaliação do <i>Test of Written English</i>	33
6 - Escala para a avaliação holística da Universidade de Tel-Aviv.....	34
7 - Escala de critérios para a avaliação da expressão escrita do <i>Certificate of Advanced English</i>	35

RESUMO

O objetivo desta pesquisa foi analisar os resultados da avaliação da expressão escrita pelo emprego de duas técnicas distintas de correção: a holística e a analítica. Suspeitando que uma das causas mais prováveis das distorções na avaliação era a ausência ou inadequação dos critérios para a atribuição das notas, vinte redações foram corrigidas por dez professores que, num primeiro momento, aplicaram a técnica holística e, posteriormente, a analítica. Com isto foi possível observar a variabilidade dos escores em cada técnica. Os resultados mostraram que há uma diferença significativa entre as notas atribuídas, quer comparando as notas de um mesmo professor face a uma e outra técnica, quer comparando os dez professores em cada uma das técnicas. Pudemos perceber que há correlação entre os escores das duas técnicas, interação dos fatores técnica utilizada, professor e redação em sua ação sobre a variável dependente (notas) e flutuações na avaliação de redações, independentemente da técnica de correção utilizada. Estes fatos nos levam a repensar o sistema de avaliação da escrita, na busca de alternativas que minimizem os efeitos da subjetividade do professor no processo.

ABSTRACT

The aim of this research was to analyze the results of two techniques for assessing writing: holistic and analytical. As we suspected that the most probable cause of distortion in the scoring process was the absence or inadequacy of scoring criteria, twenty compositions in English written by Brazilian university students were submitted to the judgement of ten teachers. The teachers were to score the compositions using first the "holistic technique" and then the "analytical technique". The results of these scorings were analyzed and it was found that there was significant disagreement among the variables. One teacher scoring the same set of compositions twice, but using different techniques, gave significantly different grades, and different teachers scoring the same set of compositions twice also presented significantly different scores. The study showed that there were correlations, interaction of the factors upon the dependent variable and fluctuations in the scoring of compositions regardless of the scoring technique used. These results suggest that we need to look closely at the way writing is assessed, and that we need to find ways of reducing the effects of subjectivity in the scoring process.

1. INTRODUÇÃO

1.1 Contexto do problema

A questão da avaliação de redações assedia de angústias o cotidiano dos professores que, muitas vezes, não conseguem sequer formalizar suas queixas e dúvidas. O fato é que as flutuações de julgamentos, provadas pelo emprego de técnicas estatísticas rigorosas, provaram as falhas, em termos de confiabilidade, na avaliação das redações. Um deslize por parte do avaliador de uma redação pode afetar a pessoa humana, pode representar uma injustiça perante o aluno. Problemas de injustiças de toda ordem decorrem desta fragilidade: por um lado, alunos menos competentes são aprovados e outros mais competentes são, por vezes, reprovados. Este tipo de erro não pode ocorrer e, na medida do possível, dever-se-á fazer de tudo para minimizá-lo. É possível inferir uma miríade de fatores que contribuem para este processo, mas isto seria muito abrangente para uma dissertação de mestrado. Desta forma, no decorrer de nosso trabalho, procuramos nos ater somente à metodologia da avaliação de redações e ao emprego de duas técnicas de avaliação de língua escrita.

1.2. Formulação do problema e sua delimitação

É evidente que nosso interesse pessoal e profissional está refletido em nosso trabalho, mas a partir do momento em que o assunto é visto sob o aspecto metodológico e as técnicas de avaliação empregadas são testadas com instrumentos estatísticos, vemos que o interesse científico também se faz presente.

Estabelecemos como objetivo fundamental do presente trabalho a análise de duas técnicas de avaliação (holística e analítica) da expressão escrita em língua estrangeira (língua inglesa), sob o aspecto de sua confiabilidade. Neste contexto o trabalho teve por objetivo verificar se a nota atribuída às redações segundo uma técnica acompanham as da outra e como agem as variáveis

independentes - técnica de avaliação, professor e redação - sobre a variável dependente - notas atribuídas: se isoladamente, ou em interação umas com as outras.

Esta restrição das variáveis independentes em detrimento de outros fatores que possam influenciar no julgamento de redações - a personalidade do avaliador, o tipo de aluno, o preparo do professor e aluno para desempenhar as respectivas tarefas, o cansaço físico ao avaliar redações, só para citar alguns exemplos - foi uma escolha feita em função dos limites de tempo, do tipo de trabalho desenvolvido e da própria metodologia da pesquisa.

1.3. Justificativa do Problema

A importância do problema apresentado não está ligada a uma proposição definitiva de soluções imediatas. A reflexão face a este problema e a possibilidade de abertura de novos caminhos para outras pesquisas constituem por si só justificativa para o nosso trabalho. Segundo Hamp-Lyons, citada em Kroll, 1990:82, não há maneira de solucionar os problemas que envolvem a avaliação de escrita enquanto não olharmos para esta atividade como um todo que envolve as seguintes partes: tarefa a ser realizada, autor do texto, procedimento de correção e o leitor do texto interagindo como numa reação em cadeia. Nossa pesquisa tenta, na medida do possível, incitar esta reflexão sobre o tema da avaliação de expressão escrita na intenção de auxiliar na busca de maneiras mais adequadas para o ensino e a avaliação da mesma.

1.4. Esquema de apresentação do trabalho

Na tentativa de alcançar o objetivo proposto, no primeiro capítulo apresentamos o conceito de escrita e uma revisão bibliográfica contextualizando a mesma no ensino de língua estrangeira. A seguir, no segundo capítulo, fazemos uma incursão no conceito "avaliação", apresentando os diferentes tipos de teste, as características essenciais de um bom teste, a redação como instrumento de avaliação de escrita, os tipos de redação e as técnicas de avaliação holística e analítica. No

terceiro capítulo apresentamos a metodologia utilizada em nosso trabalho. O capítulo quatro apresenta os resultados da nossa pesquisa, ou seja, a análise e interpretação dos dados. Finalmente, o quinto capítulo conclui nosso trabalho com considerações finais, recomendações para pesquisas futuras na área de avaliação da expressão escrita e recomendações para o ensino.

2. A ESCRITA NO CONTEXTO DO ENSINO DA LÍNGUA ESTRANGEIRA

2.1. Visão Geral

Neste capítulo procuraremos traçar um pequeno histórico sobre as abordagens do ensino da língua estrangeira que mais se destacaram a partir do século XIX. Interessa-nos mais especificamente, o papel que o ensino da língua *escrita* assumiu dentro de cada um desses métodos.

Durante muito tempo, a escrita ocupou lugar de destaque no processo de ensino - aprendizagem nas escolas. A noção de "ser educado" trazia implícita a noção de capacidade de ler e escrever na língua materna e eram estas habilidades que os alunos praticavam em sala de aula. Este enfoque dado à aprendizagem da língua materna foi transferido para o ensino de línguas estrangeiras, sem que se ponderasse sua adequação ou inadequação a uma situação em que o aluno ainda não tinha o domínio da forma falada da língua alvo. Os exercícios escritos eram uma forma fácil de manter os alunos ocupados e bem comportados. Além disso, a formulação de tais exercícios parecia ser simples e o professor podia inclusive tirá-los do livro-texto, que geralmente incluía um quadro de respostas aos exercícios. Os mesmos podiam até ser corrigidos em trabalho de grupo, fornecendo assim amplo material de avaliação para a obtenção das notas que o professor registraria em seu diário. É neste contexto geral que procuraremos, antes de mais nada, expor o conceito de escrita, para então apontar os caminhos que o ensino da escrita percorreu até os nossos dias. Por fim procuraremos situá-lo no contexto atual da metodologia do ensino de línguas.

2.2. O conceito de "escrita"

Com base nas contribuições dos autores sobre a "escrita", esta pode ser encarada sob três aspectos diferentes: a escrita como um sistema de formas convencionais que representam sons, sílabas ou palavras de uma língua, a escrita como processo criativo e a escrita como instrumento de comunicação.

A visão de escrita como um sistema de símbolos nos leva aos primórdios da lingüística e ao renovado interesse na história e na comparação destes diferentes sistemas. Tal ponto de vista entende a escrita como um sistema de formas convencionais que privilegia o texto final e relega o conteúdo do mesmo a um plano inferior ao da forma. Este modo estruturalista de conceber a escrita teve vários representantes, entre os quais podemos citar Bloomfield (1933) que foram, aos poucos, tendo suas idéias superadas pela nova concepção de "escrita" como um processo criativo.

Esta nova visão, segundo Richards, Platt e Platt (1993), é caracterizada por três etapas: a de pré-escrita, a de escrita e a de pós-escrita. A etapa de pré-escrita é aquela em que o autor organiza idéias relacionadas ao tópico a ser desenvolvido, antes de começar a escrever. Em seguida, na etapa de composição, o autor planeja, produz um esboço de suas idéias e compõe o texto. Finalmente, na etapa de pós-escrita, o autor analisa, revisa e reescreve o texto de forma a ficar a contento. Diferentemente do anterior, este enfoque processual da escrita privilegia o processo de criação de um texto e procura se ater mais ao conteúdo que à forma. Segundo Leki (1991), o método do processo encoraja os alunos a se arriscarem mais e a dividirem suas experiências com os colegas e o professor, numa tentativa de descobrir como se expressar mais efetivamente. Enfatizando esta idéia, Bowen e Marks (1994: 143) ressaltam que com este novo enfoque, a escrita deixa de ser um processo silencioso e solitário. Como expoentes desta visão de escrita podemos citar White e Arndt (1994), King (1988) e Jacobs et alii (1983).

Concluindo esta parte do capítulo, vejamos a posição de autores como Coe, Rycroft e Ernest (1994: 4) que também consideram a escrita como um processo criativo, mas vão além, definindo-a como um ato de comunicação. Este envolve muito mais que vocabulário e gramática. Seu objetivo é a transmissão de idéias por símbolos escritos. Ressaltando ainda a visão de que a escrita é um instrumento de comunicação, Brookes e Grundy (1990: 15) a definem como um modo de persuasão, do qual as pessoas se utilizam para fazer com que outras acreditem em suas idéias ou realizem suas vontades. Por sua vez, Raimes (1983: 4) afirma que a escrita é uma maneira de um autor se comunicar com o leitor, expressando suas idéias sem a presença física do outro.

Para finalizar, gostaríamos de citar White e Arndt (1994), que conseguem unir a idéia de escrita como um processo criativo com a de instrumento de comunicação. Para eles, a escrita é um registro permanente, que transcende o tempo e o espaço, sendo uma forma de expressão e um meio de comunicação. É com esta idéia globalizante de escrita que estaremos lidando no decorrer deste trabalho. De acordo com a mesma, a escrita não é apenas um sistema de convenções de uma língua, mas também o modo pelo qual a comunicação escrita se concretiza.

2.3. Evolução do ensino da escrita

2.3.1. O Método da Gramática e Tradução

O primeiro método que se destacou no ensino de línguas estrangeiras foi o **Método da Gramática e Tradução**, também conhecido nos Estados Unidos como "**Método Prussiano**", devido à sua origem germânica. Este método teve suas raízes no ensino formal do grego e latim e influenciou toda a metodologia do ensino de línguas. Tendo como principais precursores Johann Seidenstücker, Karl Plötz, H.S Ollendorf e Johann Meidinger (Rodgers, 1986:3), tal método tinha por finalidade inculcar o entendimento da gramática, expressa em termos tradicionais e treinar o aluno para escrever corretamente mediante a prática regular de versões. As aulas eram ministradas na língua materna dos alunos, com pouca utilização da língua estrangeira. Consequentemente os alunos, habituados a tal método, tinham pouco domínio da língua falada e apresentavam dificuldades para a expressão oral. Havia ênfase na aquisição do vocabulário, tradução e exercícios escritos, intermináveis e destituídos de significado, uma vez que todo texto produzido era tratado como um exercício de análise gramatical, voltado para a forma da mensagem, não para o conteúdo. Desta forma, segundo Richards e Rodgers (1986:15), os críticos mais severos deste método afirmavam que o mesmo "ensinava ao aluno tudo *sobre* a língua-alvo e não *a* língua em si".

Considerando o ensino das quatro habilidades (leitura, escrita, expressão oral e compreensão auditiva), o método em questão privilegiou o ensino de leitura e escrita, desde que estas

habilidades fossem utilizadas para uma melhor interpretação dos clássicos ou para sua tradução. Num primeiro momento, tal enfoque era inevitável, pois não havia uma metodologia específica para o ensino de línguas estrangeiras: esta deveria então ser modelada pelos métodos já empregados no ensino de uma língua antiga, no caso, o latim, língua essa já não mais usada para comunicação e cuja pronúncia original era obscura. Com o passar do tempo, entretanto, tal método se mostrou permeado por objetivos limitados cujas técnicas realmente só alcançavam tais objetivos quando o grupo de alunos era altamente intelectual e interessado em raciocínio abstrato (Rivers, 1975). Além disso, para a grande maioria dos alunos, a ênfase constante atribuída à tradução incutia a crença ingênua de que as línguas guardariam entre si uma equivalência de palavras que lhes cabia identificar, utilizando-se dos dicionários. Tal concepção, uma vez estabelecida, tornava-se um entrave à elaboração das vinculações diretas entre o significado e sua forma de expressão na língua estrangeira, processo esse fundamental para a compreensão rápida da comunicação oral e escrita, bem como para a fluência oral e a redação genuinamente clara.

Em meados do século XIX, entretanto, surgia na Europa uma demanda de comunicação entre os vários países, o que implicava na necessidade de falar outros idiomas e não apenas de ser capaz de ler e traduzir textos fora do contexto; a partir de então, a metodologia do ensino de línguas estrangeiras começou a ser remodelada.

2.3.2. O Método Direto

Nascido na França e Alemanha no final do século XIX, este foi o mais popular dos métodos naturais e teve como um de seus principais precursores Franke, com sua obra *Die praktische Spracherlernung auf Grund der Psychologie und der Physiologie der Sprache dargestellt*, publicada em 1884. Segundo os defensores deste método, as pessoas começavam a entender uma língua ouvindo-a intensamente e se tornavam proficientes na fala, falando; desta forma, aboliu-se o uso da língua materna nas salas de aula. O objetivo principal do método era desenvolver a capacidade de pensar na língua-alvo, fosse conversando, lendo ou escrevendo. As aulas de gramática

e tradução que até então eram predominantes deram lugar à prática exaustiva de frases e palavras estrangeiras correspondentes a objetos e ações na sala de aula. A gramática era ensinada através do modo indutivo e grande ênfase era dada à pronúncia correta, o que exigia professores que fossem falantes nativos ou que tivessem fluência quase que nativa para dar aulas. A tradução foi abolida por inteiro, mas, no que diz respeito ao ensino da língua escrita, usava-se a transcrição e produziam-se sumários do que fora lido, ou simples relatos do que fora discutido. Como os sons da língua estrangeira preenchiam a sala de aula o tempo todo, todas as atividades estavam ligadas ao seu uso na fala ou na escrita.

Este método provou ser eficiente na tarefa de liberar os alunos das inibições comuns ao estágio inicial da aprendizagem da língua estrangeira. Todavia, sua falha principal era a de lançar o aluno a expressar-se cedo demais na língua estrangeira, desenvolvendo, assim, uma fluência desembaraçada mas incorreta, e revestindo as estruturas da língua materna com formas da língua estrangeira.

Tal método exigia grande esforço por parte do professor, visto que era necessário que fosse fluente na língua e dotado de grande imaginação e habilidade frente à árdua tarefa de tornar o significado claro, sem recorrer em tempo algum ao uso da língua materna. Segundo Rivers (1975), o maior sucesso desse método era alcançado em situações que permitiam ao aluno ouvir e praticar a língua fora da sala de aula.

Por volta de meados dos anos 20 deste século, o uso de tal método havia decaído um pouco. Na Europa e nos Estados Unidos começavam a ser realizados estudos sobre o ensino de línguas estrangeiras nas escolas e universidades. A publicação em 1929 do trabalho do norte-americano Coleman, intitulado *The teaching of modern foreign languages in the United States*, mais conhecido como "O Relatório Coleman", evidenciou que a maioria dos alunos norte-americanos estudava uma língua estrangeira por um período de apenas dois anos. Tal relatório sustentava que o único objetivo possível e atingível neste curto espaço de tempo era o desenvolvimento da capacidade de leitura. Surgia então um novo método.

2.3.3. O Método da Leitura

Este método foi criado como resultado do Relatório Coleman. O mesmo recomendava aos professores que procurassem encontrar os meios mais eficientes para desenvolver a habilidade de leitura de seus alunos. O objetivo era fazer com que o aluno, ao terminar o curso de línguas, estivesse apto a ler na língua-alvo sem grandes problemas. Segundo Celce-Murcia (1979), tal método seria interessante para aquelas pessoas que não viajam para o exterior e não têm a necessidade de comunicar-se com falantes nativos da língua alvo; para este tipo de pessoas a habilidade da leitura seria, de fato, a mais importante.

O **Método de Leitura** sugeria que o aluno fosse treinado a ler na língua estrangeira com a identificação direta do significado, sem um esforço consciente de traduzir. A escrita devia ser limitada a exercícios destinados a fixar o vocabulário e estruturas essenciais para a compreensão do texto. Como não poderia deixar de ser, o método de leitura aumentava a capacidade de ler dos melhores alunos, mas era um obstáculo para aqueles que tinham dificuldades de leitura em sua própria língua. Segundo Rivers (1975 : 22), "o método de leitura, quase que na sua totalidade, produziu alunos que eram incapazes de compreender e falar a língua além do nível da mais simples das trocas em comunicação oral". Com o advento da Segunda Guerra Mundial, os contatos entre as nações provaram que a capacidade de leitura, por si só, não era suficiente como produto final de um curso de línguas. A partir dessa nova realidade, desenvolveu-se o método audiolingual. Na verdade, houve um despertar universal para a importância da possibilidade de falar e entender uma língua estrangeira quando usada por um falante nativo, e uma crescente impaciência com relação aos métodos que ensinavam os alunos a ler fluentemente e escrever com certa facilidade na língua estrangeira, mas deixavam os mesmos desamparados perante o fluxo normal da fala de um falante nativo.

2.3.4. O Método Audiolingual

A nova ênfase na capacidade de comunicação numa língua estrangeira levou à formação do termo "**audioral**". O mesmo servia para designar o novo método que pretendia desenvolver primeiro as habilidades da fala e compreensão da mesma, constituindo dessa forma a base para o desenvolvimento das habilidades da leitura e escrita. Entretanto, foi um professor da Universidade de Yale, Nelson Brooks (Richards e Rodgers, 1986), quem usou pela primeira vez o termo que ficaria como nome de batismo do novo método: "**audiolingual**". O uso do novo termo se deu em seu livro *Language and language learning: theory and practice*, publicado em 1964.

As origens do **Método Audiolingual** remontam ao trabalho de lingüistas estruturalistas como Bloomfield (1933) e antropólogos culturais norte-americanos como Fries (1945), que trabalhavam adotando a mesma linha de abordagem dos psicólogos behavioristas. A linha behaviorista de pesquisa enfatizava que o uso da língua era um conjunto de hábitos, estabelecidos pelo reforço ou prêmio no contexto social, ou seja, a língua materna seria um comportamento adquirido, conquistado pela criança em primeiro lugar na sua forma oral. Por isto, os teóricos da época frisavam que esta seria a melhor forma de se ensinar uma língua estrangeira: ensinando a forma oral da língua alvo, antes da forma escrita.

Segundo Moulton, citado em Rivers (1975), a aplicação dos princípios linguísticos ao ensino de línguas estrangeiras segundo o método audiolingual pode ser resumida nos cinco slogans que guiavam os professores em sala de aula:

Língua é fala, não a escrita.

Uma língua é um conjunto de hábitos.

Ensine a língua, não a respeito dela.

Uma língua é aquilo que seus falantes nativos dizem, não o que alguém acha que devam dizer.

As línguas são diferentes.

Considerando estes cinco slogans, fica óbvio que o método audiolingual tem por objetivo ensinar as habilidades lingüísticas na seguinte sequência: ouvir, falar, ler e escrever; portanto, é a linguagem oral de uso cotidiano que recebe maior atenção. Através de técnicas de repetição,

memorização e prática de padrões estruturais, o aluno recebe treinamento perfeito na produção e manipulação de elementos estruturais que ele deve aprender a usar posteriormente, sem atenção consciente para com suas características sistemáticas.

Embora não ocupem o primeiro plano, a leitura e a escrita não são negligenciadas. Todavia, não se deixa o aluno à vontade para aprendê-las à sua maneira e da melhor forma que puder, usando a língua materna como base para o pensamento. O aluno é treinado a desenvolver gradualmente sua habilidade nessas áreas (leitura e escrita), capitalizando seu crescente conhecimento da estrutura da língua. Tal processo é mantido até o ponto em que tanto a leitura como a escrita tenham se tornado para ele não meros exercícios de transposição de um idioma para o outro, mas atividades a serem conduzidas inteiramente na língua estrangeira.

Este método atingiu o seu ápice de sucesso nos Estados Unidos dos anos 60, onde era utilizado não só para ensinar línguas estrangeiras aos norte-americanos, mas também para ensinar inglês como segunda língua ou como língua estrangeira no mundo todo. A reação teórica aos princípios que norteavam o método audiolingual veio principalmente de Chomsky (1966), lingüista do Massachusetts Institute of Technology, que não só rejeitava a descrição da língua pelos estruturalistas, mas também a teoria behaviorista de aquisição de uma língua. Segundo Chomsky (1966:153), "a língua não é uma estrutura de hábitos e o comportamento lingüístico normal envolve inovação, formação de novas sentenças e padrões que estejam de acordo com regras de grande abstração e complexidade". A Teoria de Gramática Transformacional proposta por Chomsky defende que as propriedades fundamentais da língua derivam de aspectos inatos da mente e de como as pessoas processam as experiências através da língua. Tal teoria revolucionou a lingüística norte-americana da época e chamou a atenção dos lingüistas e psicólogos para as habilidades mentais que as pessoas utilizam no uso de uma língua e na sua aprendizagem. Além disso, Chomsky também propôs uma teoria alternativa sobre a aquisição de línguas, uma vez que não concordava com aquela proposta pelos behavioristas. Os behavioristas se referiam à aquisição de línguas como algo, em princípio, similar a qualquer outro tipo de aprendizagem, ou seja, a aprendizagem de línguas obedeceria às mesmas leis de estímulo e resposta, reforço e associação, como em todas as outras

coisas. Chomsky argumenta que tal teoria de aprendizagem não pode servir de modelo para a aprendizagem humana, visto que uma grande parte da linguagem humana não constitui comportamento imitado, mas é criado e repetido novamente de forma diferente daquelas regras abstratas que estão latentes ao conhecimento. Segundo Chomsky (1966), as sentenças não são aprendidas por imitação e repetição, mas são geradas a partir da competência subjacente do aprendiz. Desta forma, e de repente, todo o paradigma audiolingual foi questionado: a prática de padrões da língua alvo, bem como a repetição e a memorização. Isto gerou uma crise nos meios do ensino de línguas.

2.3.5. A Teoria Cognitiva

Surgida como um paliativo entre a reação radical anti-behaviorista da Teoria Gerativa de Chomsky e as características behavioristas do Método Audiolingual em si, esta teoria propõe que haja ênfase na competência comunicativa, sem incentivar a pronúncia perfeita, uma vez que é considerado fútil ter uma pronúncia idêntica a do falante nativo. A aquisição de uma língua é vista como um processo de formação de regras e não mais de hábitos; os erros são tidos como produtos inevitáveis do processo de ensino-aprendizagem e as quatro habilidades têm o mesmo grau de importância. Sempre se procura uma contextualização de todos os pontos para que o aluno possa aprender através do significado. Esta prática se baseia numa visão de aprendizagem que permite uma concentração consciente na gramática e reconhece a função dos processos mentais abstratos na aprendizagem, e não define o processo de aquisição de uma língua meramente como uma questão de formação de hábitos.

Segundo Richards e Rodgers (1986), houve na década de 70 um considerável interesse por esse tipo de trabalho com línguas estrangeiras, mas nenhuma linha de pesquisa bem definida emergiu,

nem mesmo um único método utilizando as premissas acima citadas. A falta de uma alternativa para o Método Audiolingual nos Estados Unidos tem levado a uma situação de adaptação, inovação, experimentação e muita confusão. De um lado, segundo Richards e Rodgers (1986), há novos métodos que se desenvolveram à margem das teorias linguísticas e de aquisição da segunda língua, como por exemplo: "Total Physical Response", "The Silent Way" e "Counseling-Learning". De outro lado, há abordagens que dizem derivar de teorias contemporâneas da linguagem, como por exemplo o Método Natural e a Abordagem Comunicativa. Entre todas estas novas tendências, a que mais nos interessa, por acreditarmos ser a mais completa de todas, é a abordagem comunicativa, cujos princípios norteadores e as principais características apresentaremos a seguir.

2.3.6. A Abordagem Comunicativa

Surgida no final da década de sessenta, a **Abordagem Comunicativa** tem sua origem na insatisfação de um grupo de lingüistas ingleses em relação aos métodos de ensino de línguas até então em vigor. Dentre estes autores, podemos citar lingüistas como Firth (1957) e Halliday (1970), que viam a necessidade de centrar o ensino de línguas estrangeiras na proficiência comunicativa e não no mero domínio de estruturas linguísticas. Segundo Richards e Rodgers (1986), este ponto de vista teve o respaldo dos teóricos Candlin (1976) e Widdowson (1972), dos filósofos Austin (1978) e Searle (1969) e dos sociolingüistas norte-americanos Hymes (1972), Gumperz (1972) e Labov (1970). Os mesmos acreditavam que o desenvolvimento da competência comunicativa dos falantes era o principal objetivo do ensino de línguas. Esta visão da linguagem é considerada pela maioria dos autores uma abordagem e não um método, pois permite muito mais flexibilidade e variedade na sua interpretação do que um método permitiria.

O princípios da **Abordagem Comunicativa**, segundo Richards e Rodgers (1986), são:

1. A língua é um sistema para a expressão do significado.
2. A função básica da língua é a interação e a comunicação.

3. A estrutura de uma língua reflete seu uso funcional e comunicativo.
4. As unidades primárias de uma língua não são características meramente gramaticais e estruturais, mas categorias de significado funcional e comunicativo, como exemplificado no discurso do falante.

A teoria de aprendizagem de uma língua que se baseia na **Abordagem Comunicativa** enfatiza as maneiras de desenvolver as habilidades comunicativas, e as atividades em sala de aula devem enfatizar tarefas passíveis de realização através do uso da língua, ou que envolvam negociação ou interação para a obtenção de informações. A adoção de tal abordagem para o desenvolvimento das habilidades linguísticas faz com que o professor dê igual crédito ao ensino da escrita e da leitura de um lado, e ao ensino de produção oral e compreensão auditiva de outro. Conseqüentemente, tanto o ensino da leitura como da escrita pode ter início já no primeiro dia de aula, como afirmam Finocchiaro e Brumfit (1983).

No que diz respeito ao ensino da escrita, especificamente, a abordagem comunicativa enfatiza o fim para o qual um texto é elaborado e a sua audiência, e o aluno é encorajado a escrever como um escritor de verdade, tendo sempre em mente as duas questões principais:

- Por que eu estou escrevendo este texto?
- Quem vai ler este texto?

Tradicionalmente o professor, e só ele, tem sido a audiência do texto produzido pelo aluno. Entretanto, teóricos como Finocchiaro e Brumfit (1983) e White e Arndt (1994) acreditam que os escritores escrevem melhor quando envolvidos em um ato verdadeiramente comunicativo, ou seja, quando um escritor escreve para um leitor de verdade. Assim sendo, os professores envolvidos na **Abordagem Comunicativa** têm dado textos produzidos por alunos a outros alunos, os quais não apenas os lêem, mas os reescrevem de outra forma, os resumem ou fazem comentários, sem no entanto corrigi-los. Além disso, o professor pode especificar uma audiência fora da sala de aula, fornecendo desta forma aos alunos o contexto para o qual deverão produzir seus textos, bem como o grau de formalidade e o conteúdo a ser abordado. Na **Abordagem Comunicativa** os alunos podem representar personagens para as quais cartas foram escritas, trocar cartas entre si ou entre turmas diferentes, sempre respondendo, fazendo comentários ou esclarecendo dúvidas, sem jamais corrigir o

texto do colega. Apesar de toda essa tentativa de deslocar o enfoque do domínio das estruturas para a obtenção de uma competência comunicativa, para muitos professores "a **Abordagem Comunicativa** ainda mantém a concentração no produto e não no processo de escrita"(Raimes: 1983). Esta mudança de enfoque vai se dar por volta do início da década de oitenta, quando surge a **Abordagem do Processo** (*Process Approach*), segundo o qual a atenção se desloca do produto da escrita para o processo em si.

2.3.7. A Abordagem do Processo

Nomes como Elbow (1973), Graves (1983), Flower e Hayes Junior (1981), Zamel (1982), Raimes (1983), Bereiter e Scardamalia (1987), White e Arndt (1994) vão se destacar nesse novo panorama do ensino de língua escrita, porém ainda dentro da Abordagem Comunicativa.

Segundo a **Abordagem do Processo**, além do propósito da elaboração do texto e para quem ele deve ser escrito, outras duas perguntas passam a ter importância para o aluno:

- Como eu escrevo isto?
- Como é que eu começo?

Dentro desta perspectiva, todos os escritores tomam decisões sobre a maneira de começar e organizar um texto. Os alunos, em particular, precisam se acostumar com a idéia de que, o que eles colocam no papel nem sempre é o produto final de seu raciocínio, mas sim o começo, um esboço. Eles não devem esperar que as primeiras palavras escritas sejam perfeitas para aquela situação, e precisam se habituar com a idéia de reler e reescrever um mesmo texto várias vezes. Segundo estes teóricos, quando se dá a um aluno o tempo para processar seu texto e ter o "feedback", tanto do professor quanto dos colegas, sobre o que escreveu, ele descobrirá novas idéias, novas frases, novas palavras e, na medida em que for reestruturando seu texto , produzirá novos esboços, até chegar ao texto final.

Desta forma, na **Abordagem Processual** os alunos não escrevem sobre um certo tema dado em um tempo limitado e entregam a composição ao professor para que seja "corrigida" - o que

geralmente significa "encontrar os erros". Ao contrário, os alunos exploram um tópico através da escrita, mostrando seus esboços ao professor e aos colegas e utilizando os comentários e sugestões para relerem seus textos, reestruturarem suas frases, incorporar novas idéias, eliminarem redundâncias, ou seja, revisarem seus textos.

Os professores que adotam tal procedimento dão a seus alunos dois suportes básicos e cruciais para o processo da escrita: *tempo* para os mesmos testarem suas idéias e "*feedback*" quanto ao conteúdo dos seus esboços. Segundo Taylor, citado em Raimés (1983 : 11), esses professores "acabam por concluir que, desta forma, o processo de escrita se torna um processo de descoberta para os alunos: descoberta de novas idéias e de novas formas de expressá-las".

2.4. Considerações Finais

A título de conclusão deste capítulo gostaríamos de deixar bem claro que tanto os métodos do ensino de língua estrangeira e as teorias lingüísticas das quais emanam, quanto as abordagens praticadas com maior ou menor ênfase em cada um dos períodos aqui descritos, atualmente se sobrepõem. Isto acontece em decorrência da existência de uma grande variedade de professores de línguas, de diferentes estilos de aprendizagem, de diferentes traços psicológicos de cada aluno e de cada professor . Conseqüentemente, não há um único método de ensino que seja melhor que os outros, assim como não há uma única maneira de ensinar. No que concerne ao ensino da língua escrita, que é o nosso interesse principal, acreditamos que a **Abordagem de Processo** é a mais adequada para o ensino de escrita em uma segunda língua, uma vez que prima pelo processo e não pelo produto final, embora saibamos que isto dificulta a avaliação, já que a mesma necessariamente se realiza através do julgamento de um produto final.

3. A AVALIAÇÃO DA EXPRESSÃO ESCRITA

3.1. Visão Geral

Neste capítulo apresentaremos os diferentes tipos de teste, as características essenciais de um bom teste, o conceito de redação como instrumento de avaliação de escrita, os tipos de redação e as técnicas holística e analítica de avaliação de redação.

3.2. Tipos de teste

Um teste é um instrumento para obtermos informações (Hughes, 1994 : 9) e estas informações variam de situação para situação. Em termos gerais e de acordo com a nomenclatura da bibliografia da área, há vários tipos de teste de língua, dentre os quais ressaltaremos os testes denominados "discrete point", os integrativos, os comunicativos, os diretos e os indiretos.

3.2.1. Testes "discrete point "

Um teste denominado *discrete point* é aquele que se refere à avaliação de apenas um elemento da língua por vez, item por item. Estes, segundo Hughes (1994 : 19), são os que servem de instrumento para medir o grau de domínio de um elemento de uma determinada língua, num determinado momento. Os mesmos se baseiam na teoria de que a língua é composta de diferentes partes e que uma vez que o aluno seja capaz de usar estas diferentes partes adequadamente, ele domina aquele item específico. Por exemplo, se um aluno entende e consegue utilizar a terminação "s" da terceira pessoa do singular do presente do indicativo do inglês, então ele domina o "presente simples" deste idioma. Utilizando-se este tipo de teste, elementos como tempos verbais, preposições, artigos, sufixação e prefixação, entre outros, podem ser testados separadamente.

O *discrete point test* tem suas vantagens e desvantagens. Segundo Weir (1990 : 2), a maior vantagem é a sua objetividade, que o torna mais confiável, pois os itens são facilmente quantificáveis e permitem uma maior abrangência. Embora tais vantagens possam ser atraentes para muitos

professores e elaboradores de testes, não o são para outros. Os descontentes com este tipo de avaliação objetiva encontram apoio em teóricos como Oller, que preconizam a ineficiência deste tipo de avaliação, pelo fato de separar os elementos da língua em partes e fazendo com que as propriedades da mesma língua se percam com esta separação. Segundo Oller (1986) a língua, como todo sistema, é formada por partes que interagem, dando origem às suas características típicas. Por isso mesmo não pode ser fracionada, nem no momento do ensino, nem no da avaliação da aprendizagem. Para Oller o principal problema deste tipo de teste é sua inerente insuficiência como indicador de proficiência na língua, uma vez que considera apenas fragmentos da referida língua e não o processo de comunicação como um todo. Esta insatisfação gerada pelo uso indiscriminado de testes deste tipo fez com que novas propostas de avaliação surgissem, como os chamados *testes integrativos*.

3.2.2. O Teste Integrativo

Ao contrário dos "discrete point tests", a avaliação integrativa exige que o aluno seja capaz de combinar vários elementos da língua para realizar uma tarefa. Esta pode envolver a elaboração de uma redação, a tomada de notas durante uma palestra, um ditado ou o "cloze". Segundo Oller (1986 : 37), os testes integrativos "são uma tentativa de avaliar a capacidade do aluno de utilizar muitos elementos da língua ao mesmo tempo, aproximando, desta forma, o aluno do processo real de comunicação e utilização da língua". Alguns procedimentos de avaliação, tais como as entrevistas orais e as redações, são por ele consideradas como "testes integrativos globais", uma vez que vão além da mensuração de uma parte limitada da competência linguística.

Entretanto, nem todos advogam a mesma causa de Oller. Alderson (1980), por exemplo, levanta a questão da validade do teste integrativo, exclusivamente no que diz respeito ao "cloze" como forma de avaliação. As controvérsias geradas pelo confronto direto entre o teste objetivo e o integrativo propiciaram o surgimento de novas abordagens da avaliação.

3.2.3. Teste direto *versus* indireto

A dicotomia *teste integrativo* versus teste "*discrete point*" está diretamente relacionada a uma outra idéia de oposição, a do teste direto/indireto. Um teste integrativo é, em geral, um *teste direto*; ou seja, exige que o candidato desempenhe exatamente a habilidade que se deseja medir. Por exemplo, se se quer testar a capacidade de comunicação oral de um candidato, faz-se com que fale. Se desejamos verificar a capacidade de expressão escrita de alguém, faz-se com que esta pessoa escreva. As atividades ou tarefas a serem realizadas devem, contudo, ser as mais autênticas possíveis. Um teste "*discrete point*", por sua vez, quase sempre é um *teste indireto*. Por indireto entendemos todo tipo de teste que mede habilidades subjacentes às quatro habilidades básicas (fala, escrita, compreensão auditiva e leitura) e não as habilidades em si. Como exemplo podemos citar o método proposto por Lado (1961) para testar a pronúncia. Nele o candidato tinha que ler uma lista de palavras e identificar, por escrito, pares destas palavras que rimavam entre si. Apesar desta relação ser quase sempre verdadeira, é possível que um teste integrativo seja indireto, como é caso do "cloze test", que requer um processamento do discurso em diferentes níveis e, mesmo assim, não pode ser considerado como um teste de cada nível específico. Vejamos então, a seguir, o teste comunicativo.

3.2.4. O Teste Comunicativo

Nascido da Abordagem Comunicativa do ensino de línguas, este tipo de teste enfatiza a dimensão comunicativa da avaliação, não mais a dimensão lingüística. O foco principal das atenções é o modo como o aluno se utiliza efetivamente da língua estrangeira para realizar tarefas distintas e específicas em situações contextualizadas. Para teóricos como Morrow (1977); Canale e Swain (1980 : 1-47), o teste comunicativo, além de se preocupar com o **que** o aluno sabe sobre a **forma** da língua e sobre seu uso adequado, também oferece ao aluno a oportunidade de demonstrar o que ele sabe sobre uma situação comunicativa específica, ou seja, que ele demonstre sua **habilidade de se comunicar** com fluência e adequação dentro de um determinado **contexto** sociolingüístico.

Este raciocínio teve grande receptividade entre os professores de línguas durante a década de 80, ao ponto de autores como Bachman, citado em Weir (1990: 7), caracterizarem as décadas de 80 e 90 do nosso século como as "décadas do teste comunicativo". A principal vantagem deste tipo de teste é que ele mede várias habilidades e sub-habilidades ao mesmo tempo, combinando-as entre si, seja na forma oral ou escrita. Segundo Madsen (1983), um teste comunicativo é aquele que indica quão bem e efetivamente uma pessoa se comunica em uma segunda língua.

Hughes (1994) é de opinião que um teste comunicativo deva ser aplicado quando há o desejo de medir a habilidade que uma pessoa tem para participar em atos de comunicação, sejam estes orais ou escritos, incluindo-se aí também a leitura e a compreensão auditiva.

Contudo, há restrições a este tipo de teste. O maior de todos os problemas é que, se os professores não forem bem treinados para corrigi-los, fica difícil provar que tal teste daria os mesmos resultados se aplicado em um outro dia; se não for bem elaborado, fica também difícil provar que apresentaria os mesmos resultados se aplicado a outros alunos com características semelhantes. Além disso, toda a abordagem comunicativa pressupõe que se trabalhe, na medida do possível, com situações de comunicação autênticas; mas, como podemos falar de autenticidade sob a ótica de um teste? A própria situação da sala de aula e da avaliação por si só já depõem contra a "naturalidade" do processo comunicativo. É difícil imaginar um aluno à vontade imaginando que está conversando com um amigo ou relatando suas férias de verão para seu professor; esta situação "forçada" não é nada natural. Com certeza, o aluno submetido a um teste deste tipo não será fiel à escolha de palavras usadas no relato, nos seus movimentos corporais como expressão visual, posição das mãos, etc. Contudo, o teste comunicativo sempre apresenta uma boa repercussão entre alunos e professores.

Como pudemos perceber, todos os tipos de teste apresentados apresentam ressalvas quanto à sua utilização. Cabe ao professor decidir qual teste escolher de acordo com o tipo de informação que deseja obter. Na seção seguinte procuraremos apresentar as três características essenciais de um teste.

3.3. Características essenciais de um bom teste

3.3.1. Visão Geral

Um bom teste deve ser prático, confiável nos resultados e válido em relação ao seu uso potencial futuro. Por exemplo, se um professor quer verificar a habilidade de seus alunos na leitura e compreensão de textos, ele precisa verificar se o teste é realizável, se os resultados serão consistentes e se o teste que ele estará utilizando vai medir exata e especificamente a habilidade de leitura dos alunos. Caso contrário, sua avaliação pode ficar comprometida. Vejamos então como são definidas estas características, de modo específico.

3.3.2. Praticidade

Para ser prático, um teste deve ser passível de elaboração, aplicação e correção dentro dos limites de tempo e recursos disponíveis para tal. É possível, no entanto, que um teste seja prático em sua elaboração, mas não em sua avaliação. Um exemplo disto seria uma entrevista gravada com aproximadamente trinta minutos de duração, com cada um dos vinte alunos de uma turma, havendo apenas um professor para a análise. Por outro lado, um exemplo de um teste prático na correção seria o Vestibular, corrigido através do computador; contudo, é muito difícil elaborar este tipo de teste. Como podemos observar, a questão da praticidade é relativa, pois um teste prático na elaboração nem sempre o é na correção e vice-versa. Tudo dependerá do tipo de informação que se procura obter e dos recursos disponíveis para tal.

3.3.3. Confiabilidade

Um teste é confiável se seus resultados são consistentes ou estáveis. Na avaliação de habilidades linguísticas a confiabilidade dos testes tem sido estimada de três modos diferentes: pelo

"test-retest", pela "aplicação de um teste de formas equivalentes" e pela verificação da "consistência interna".

A confiabilidade pelo "test-retest" é normalmente auferida pela submissão do mesmo grupo de alunos ao mesmo teste, por duas vezes consecutivas. A seguir, calcula-se o coeficiente de correlação entre os pares de escores das duas avaliações. Quanto mais próximos forem os escores, maior será o coeficiente e mais confiável será o teste.

O segundo modo de verificação da confiabilidade, comum em avaliação do domínio das habilidades lingüísticas, é a aplicação de **formas equivalentes de um teste** : são aplicados dois testes equivalentes entre si (pode ser , por exemplo, a versão A e a B de um mesmo teste) a um mesmo grupo de alunos e, a partir dos resultados, calcula-se o coeficiente de correlação entre os escores das duas versões. Novamente, quanto mais próximos forem os escores, maior será o coeficiente de confiabilidade do teste.

Finalmente, o terceiro modo mais comum de verificarmos a confiabilidade numa avaliação de língua é a **consistência interna**. Segundo Brown (1993 : 99), há várias maneiras de verificação deste tipo de confiabilidade, mas o mais fácil é o método da **divisão de um teste em metades** no qual (1) os itens pares e os itens ímpares do teste são corrigidos separadamente; (2) o coeficiente de correlação entre estes dois "sub-testes" é calculado; (3) o coeficiente de correlação resultante é ajustado para a confiabilidade do teste inteiro usando testes estatísticos específicos para este fim, tais como o teste alfa de Cronbach (Brown, 1993), o K-R20 de Kuder-Richardson (Montgomery, 1991) e o teste de Spearman -Brown (Brown , 1993). Este tipo de verificação de confiabilidade é preferido aos demais por ter a vantagem de utilizar apenas um teste, que é aplicado apenas uma vez, a um mesmo grupo de alunos.

Além destes três modos mais comuns de verificação de confiabilidade, há ainda outros dois que nos interessam mais, por estarem diretamente ligados à avaliação de redações: a "confiabilidade inter-avaliadores" e a "confiabilidade intra-avaliadores". Estes modos de verificação da confiabilidade normalmente são usados quando os professores precisam fazer julgamentos sobre a produção oral ou escrita de alunos (entrevistas ou redações).

A confiabilidade **inter-avaliadores** é, na sua essência, uma variação do tipo das **formas equivalentes**. Após a avaliação de um conjunto de redações por dois avaliadores, é calculado um coeficiente de correlação entre as duas séries de escores. Por sua vez, a confiabilidade **intra-avaliadores** está mais próxima do tipo "**test-retest**", pois um mesmo avaliador julga um conjunto de redações por duas vezes consecutivas e em momentos diferentes. Em seguida é calculado o coeficiente de correlação entre os escores das duas avaliações.

3.3.4. Validade

A praticidade e a confiabilidade são requisitos básicos de um bom teste, mas não são os únicos. A validade de um teste também é uma característica importante a ser considerada no momento da escolha ou adoção de um tipo de teste. Segundo Brown (1993: 101) um teste é válido se "mede exatamente o que diz estar medindo".

Alguns autores, como Hughes (1994), subdividem o conceito de validade em três partes: "validade de conteúdo", "validade de critério" e "validade de conceito". A "validade de conteúdo" responde à pergunta *Este teste é um bom exemplo do conteúdo que o mesmo quer avaliar?* A "validade de critérios" é aquela em que um teste é comparado ou correlacionado a uma medida de critério externa. Por exemplo, um mesmo grupo de alunos é submetido a dois testes diferentes: um em desenvolvimento e o outro já estabelecido e reconhecido como uma boa forma de avaliação naquela área específica. Num segundo momento, é calculado o coeficiente de correlação entre os dois grupos de escores e determinado o grau de correlação ou sobreposição entre eles. Finalmente, a "validade de conceito" se baseia no grau em que os itens de um teste refletem os aspectos essenciais da teoria em que o referido teste se baseia. Por exemplo: se desejamos testar a competência comunicativa escrita de um aluno, devemos fazer com que ele produza textos inteiros, com um objetivo comunicativo, não apenas frases soltas e desconexas.

Em contrapartida, Bachman (1991 : 242 - 243) refere-se à validade como uma medida unitária, que pode ser o resultado da aplicação de diferentes instrumentos de avaliação (testes para

medir a validade de conteúdo, conceito ou critério; por exemplo), sobre os quais são feitas inferências. O autor não discute a validade dos testes em si, mas a de seus resultados em relação ao tipo de informação que se pretende obter através dos mesmos, o objetivo do teste.

3.4. Conceito de redação como instrumento de avaliação da expressão escrita

A redação é um dos instrumentos habitualmente usados nas escolas para avaliar a capacidade de expressão escrita dos alunos. Isto ocorre porque a redação é um tipo de teste direto, em que se pede ao aluno que escreva quando se quer avaliar a sua capacidade para tal. É neste contexto que situamos a redação.

Além de ser um instrumento para medir o grau de sucesso no processo de ensino-aprendizagem, na esfera escolar, a redação é também usada por instituições de ensino como instrumento de seleção e admissão de candidatos em cursos a nível de segundo grau, concurso vestibular e cursos de pós-graduação. Assim sendo, a avaliação de redações constitui um desafio para a maioria dos professores que conhecem o significado que a expressão escrita assume no contexto social, sobretudo no escolar.

Para Soares e Campos (1979 : 85-86)

"ao escrever sobre um determinado assunto, escrevemos para atingir um objetivo, que deve estar presente em todo o texto: nas frases de introdução, desenvolvimento e conclusão, permitindo que se construa uma unidade de composição escrita - uma unidade articulada de forma articulada ou integrada. A esta unidade de composição escrita a respeito de um assunto particular produzida para atingir um objetivo, e estruturada por um conjunto de orações dá-se o nome de parágrafo e uma redação pode constituir-se de um ou mais parágrafos".

Seguindo esta linha de raciocínio Richards, Platt e Platt (1993 : 71), por sua vez, definem redação como um texto produzido para a prática de habilidades e técnicas de escrita ou para a demonstração da capacidade de escrever por parte de uma pessoa. Esta posição reflete bem a realidade da sala de aula, onde o aluno escreve apenas para provar ao professor que é capaz de

realizar a tarefa pedida, sem se preocupar com a função comunicativa do texto a ser produzido. A seguir veremos os tipos de redação e aquele que mais nos interessa neste trabalho.

3.5. Tipos de redação

As redações podem ser classificadas em um contínuo que em seus extremos apresenta as seguintes categorias: livre e controlada. Segundo Richards, Platt e Platt (1993 : 71) *redação livre* é aquela em que o aluno de escrita não é controlado de maneira alguma, ao passo que *redação controlada* é aquela em que o aluno é orientado e guiado de várias maneiras para produzir seu texto, seja através de perguntas a serem respondidas, de fotos a serem descritas ou de frases a serem completadas.

A seguir ater-nos-emos , mais especificamente, à redação acadêmica. Esta é a redação que o aluno de um curso de graduação universitária precisa elaborar no decorrer de sua formação para desenvolver sua capacidade de expressão escrita. É difícil enquadrar a redação acadêmica nas categorias "livre" ou "controlada", pois a mesma apresenta características de ambas.

Segundo Hamp-Lyons (1989 : 5) um texto classificado como redação acadêmica tem características próprias a serem consideradas, a saber:

_ é baseado em informações, ou seja, mesmo que o texto expresse opiniões, elas são baseadas em estudos cuidadosos e considerando-se as evidências sobre o assunto discutido;

_ é organizado de modo formal e previsível . O leitor de qualquer texto acadêmico espera encontrar uma introdução, um desenvolvimento com argumentação clara e organizada e uma conclusão. À primeira vista isto parece tornar a redação acadêmica complexa mas, na verdade, a torna mais simples, porque estes aspectos previsíveis são passíveis de aprendizagem e de verificação. Portanto, não é totalmente livre;

_ apresenta uso de vocabulário adequado. Qualquer assunto acadêmico tem um certo número de palavras-chave que são imprescindíveis para a redação de um texto em sua área específica (jargão). Escrever sobre tais assuntos sem dominar esse vocabulário básico mínimo pode ser

infrutífero. Para se obter um bom vocabulário, em termos gerais, é preciso fazer muita leitura sobre os temas de uma área de interesse.

Ainda discorrendo sobre a redação acadêmica, Campbell (In.: Kroll, 1990 : 211) frisa que a elaboração de uma redação acadêmica envolve processos complexos como leitura, compreensão, relatórios, planejamento, escrita, revisão e coordenação de todo o processo. Além disso, a autora menciona também que até a mais original das redações acadêmicas pode se valer de citações, paráfrases, resumos e breves referências a outros textos.

Hamp-Lyons (In.: Kroll, 1990 : 73) mostra que é importante dar condições a uma pessoa para que ela produza um texto que demonstre toda a sua capacidade de desempenho. Isto é feito, para fins de avaliação da escrita, através da manipulação de algumas variáveis nas tarefas designadas aos alunos. As variáveis a serem consideradas são: o tópico da tarefa, o objetivo do texto a ser criado, o modo do discurso a ser seguido, as expectativas culturais que envolvem a atividade (o que o aluno conhece a respeito do assunto a ser desenvolvido e sua visão sobre seu nível de desempenho na segunda língua), o público para quem se escreve, a especificação retórica da atividade (como se expressar correta e efetivamente em relação ao tópico, ao público-alvo e ao objetivo da atividade) e o nível de dificuldade lingüística envolvido na tarefa. Segundo a autora, uma vez que o aluno tenha conhecimento dessas variáveis, fica muito mais fácil organizar seu texto de modo a cumprir satisfatoriamente a tarefa de escrever.

Nossa escolha pela redação acadêmica se justifica justamente por se esperar do aluno o conhecimento prévio de algumas destas variáveis supracitadas (público-alvo, por exemplo) e o domínio das técnicas de elaboração da redação, o que permite ao professor uma melhor elaboração e aplicação de técnicas de avaliação da expressão escrita. Em outras palavras, o aluno tem conhecimento da estrutura que seu texto deve ter e o professor sabe o que deve procurar nas redações a serem avaliadas. Tudo isto faz com que a redação acadêmica fique situada num ponto entre as extremidades *livre e controlada* do contínuo "tipos de redação", aproximando-se mais das redações do tipo controlado, visto que as oportunidades para a elaboração de uma redação totalmente livre são raras. Geralmente, os alunos não podem deliberar sobre o estilo do texto a ser produzido ou o assunto

a ser abordado. Em uma situação de avaliação, então, a possibilidade de se produzir uma redação totalmente livre não existe; ela só é possível no contexto de sala de aula para promover o desenvolvimento da expressão escrita.

Na próxima parte do capítulo estaremos apresentando as duas técnicas mais comuns de avaliação de redação.

3.6. Técnicas de avaliação de redação

3.6.1. Visão Geral

As várias maneiras de se corrigir uma redação podem ser enquadradas num contínuo que apresenta como extremos duas formas de avaliação: a holística e a analítica. Genericamente falando, a última envolve a verificação detalhada de vários aspectos do texto produzido pelo aluno e a primeira consiste de uma leitura geral do texto, seguida da atribuição de uma nota. A seguir veremos as características específicas e a conceituação de cada uma.

3.6.2. A Técnica Analítica

A técnica analítica constitui um procedimento de avaliação que se baseia na análise das partes que compõem a expressão escrita ou falada. Segundo Hughes (1994 : 94), tal técnica de avaliação apresenta uma série de vantagens. Primeiramente, descarta de imediato o problema das diferenças de níveis no desenvolvimento das sub-habilidades entre os alunos, pois enfoca o texto em partes, analisando item por item; desta forma, o aluno que escreve com certo estilo, mas apresenta sérios problemas de ortografia, terá seu texto analisado sob todos os aspectos e seu texto não será subestimado por estar permeado de erros ortográficos. O contrário também é possível: um aluno pode escrever quase que sem erros ortográficos, mas com poucas idéias interessantes ou com um conteúdo que deixa a desejar. Todavia, todos estes aspectos serão devidamente analisados e ponderados. Num

segundo momento, a técnica analítica faz com que o avaliador considere aspectos do desempenho do aluno que poderiam passar despercebidos num outro tipo de avaliação. Finalmente, o simples fato de os avaliadores terem que dar notas independentes - por itens que compõem o todo - já faz com que o escore final seja mais confiável, uma vez que o avaliador tem a oportunidade de avaliar o produto final várias vezes. Entretanto, é preciso reconhecer a dificuldade que os avaliadores podem encontrar ao atribuir notas específicas para as partes que compõem o todo.

Uma outra característica essencial deste tipo de avaliação é a atribuição de pesos aos diferentes aspectos que devem ser considerados pelos avaliadores. Esta atribuição de pesos aos componentes do todo é a expressão da importância dada a cada um dos componentes pelos idealizadores dos critérios, sejam eles professores ou instituições. Em outras palavras, se a pessoa que elabora os critérios de avaliação considera tudo importante na elaboração de um texto, provavelmente cada componente terá peso igual aos demais. Por outro lado, se o autor dos critérios que devem ser adotados por um determinado grupo de avaliadores privilegia a forma em detrimento do conteúdo, provavelmente teremos uma série de critérios com pesos maiores para ortografia, organização de parágrafos e pontuação, e pesos menores para a qualidade comunicativa do texto ou da argumentação. Em qualquer um dos casos, a nota final do aluno será a soma dos escores atribuídos às partes, multiplicados pelos respectivos pesos, ou a soma de partes diferentes.

Como exemplos deste tipo de avaliação podemos citar os testes de língua inglesa elaborados pela Universidade de Cambridge, Inglaterra, que são aplicados no mundo inteiro e trazem consigo uma grande carga de tradição e um alto grau de confiabilidade, pois seus elaboradores e avaliadores são altamente capacitados e exaustivamente treinados. É importante ressaltar aqui que esses critérios não são puramente analíticos e sim uma mistura de holísticos com analíticos, uma vez que os alunos têm seus textos classificados de acordo com critérios explícitos e a interpretação dos quesitos que compõem os textos para cada uma destes níveis traz uma certa subjetividade do avaliador. Para minimizar esta nuance subjetiva na aplicação de critérios rígidos e pré-estabelecidos, a Universidade de Cambridge treina intensivamente os seus avaliadores para este tipo de teste.

Vejamos então o exemplo dos critérios adotados para a avaliação de redações que fazem parte do Primeiro Certificado de Língua Inglesa da Universidade de Cambridge (*First Certificate in English*):

Quadro 01: Critérios de avaliação do *First Certificate in English*

Nível	Uso da Língua	Tarefa
18-20	Inglês natural com um mínimo de erros.	Realização total da tarefa.
16-17	Mais que uma coletânea de frases, apresenta bom vocabulário e estrutura. Alguns lapsos eventuais.	
11-15	Naturalidade suficiente do inglês , sem apresentar muitos erros.	Realização simples, mas correta da tarefa..
8-10	Comunicação estabelecida apesar de alguns erros.	Tarefa razoavelmente realizada.
5-7	Falta de controle/ erros básicos frequentes.	Tarefa apenas parcialmente realizada.
1-4	Incoerente.	Irrelevância grosseira e/ou texto muito curto para a avaliação do uso da língua.

Fonte: Cambridge First Certificate: Examination Practice 2. Teacher's Book. Cambridge : CUP, 1986, p. 24.

Com base nos critérios acima, o professor deve ler a redação e a ela atribuir uma nota de 1 a 20, sendo que parte desta nota é reservada ao uso da língua e a outra parte à tarefa proposta. A nota final do aluno é a soma das duas partes específicas passíveis de avaliação para este nível, *Uso da Língua* e *Tarefa*, que classifica o aluno num dos níveis da primeira coluna, ou seja, se uma redação obteve a nota final igual a 14, é enquadrada na categoria entre 11-15. Lembramos ainda que esta nota final é a média das notas atribuídas a uma redação, quando avaliada por avaliadores bem treinados para este fim. Vejamos ainda da Universidade de Cambridge o exemplo de critérios de avaliação para a prova de redação do Certificado de Proficiência (*Certificate of Proficiency in English*):

Quadro 02: Critérios de avaliação do *Certificate of Proficiency in English*

Nível	Uso da Língua	Conteúdo
16-20	Ambicioso no procedimento com uso da língua de alto nível. Lapsos eventuais de um falante nativo.	Realização completa da tarefa.
11-15	Natural e apropriado no estilo, com erros eventuais. Alguma sofisticação no uso da língua.	
8-10	Estrutura e uso da língua claros, mas limitados.	Tarefa razoavelmente cumprida.
5-7	Falta de controle/muitos erros.	O assunto não foi exposto em sua devida extensão ou não adequadamente explorado.
1-4	Erros e pobreza de vocabulário prejudicam a comunicação..	Irrelevância grosseira e/ou texto muito curto para ser avaliado.

Fonte: Cambridge Proficiency Examination Practice 4: Teacher's Book. Cambridge : CUP, 1991, p. 9.

Como no exemplo anterior, a nota final atribuída a uma redação é a soma dos dois itens constitutivos, *Uso da língua* e *Conteúdo*, e a média entre diferentes avaliadores.

Um outro exemplo de avaliação analítica é o desenvolvido e aplicado pela Universidade de Edimburgo, com o apoio do *British Council*, chamado de *PMP2* (New Profile Scale and Profile Method 2), que avalia as redações com base em cinco critérios específicos: Qualidade Comunicativa, Organização, Argumentação, Acuidade Lingüística e Adequação Lingüística (ver Anexo 3).

Para não ficarmos apenas com exemplos importados de universidades estrangeiras, vejamos ainda o conjunto de critérios para a técnica analítica, utilizado por Avosani, em 1981, em sua pesquisa realizada em escolas da região metropolitana de Curitiba. O referido autor, para a obtenção do seu grau de mestre em Educação, comparou técnicas de avaliação de redações de estudantes a nível de segundo grau, tendo usado, para a técnica analítica, o seguinte conjunto de critérios:

Quadro 03: Critérios para a correção analítica, de Avosani (1981)

	Categorias	Valor
01	Apresentação geral: limpeza, margens, etc.	01
02	Emprego da crase	01
03	Ortografia	07
04	Pontuação	03
05	Flexão de substantivo e adjetivo (gen., núm., grau)	03
06	Grafia dos verbos	02
07	Partição de sílabas em finais de linha	02
08	Pronomes relativos	02
09	Pronome "se"	02
10	Emprego do infinitivo	02
11	Emprego de conectivos	03
12	Correlação de tempos verbais	02
13	Concordância (geral)	05
14	Regência (geral)	05
15	Colocação (geral)	05
16	Composição (introdução, desenvolvimento, conclusão)	05
17	Paragrafação (divisão segundo assuntos)	05
18	Idéias dos parágrafos em relação à idéia central	05
19	Emprego de idéias desnecessárias	05
20	Repetição de palavras ou expressões	05
21	Estrutura do período (sem principal, etc.)	05
22	Adequação do vocabulário às idéias	05
23	Pobreza de idéias	05
24	Concatenação lógica das idéias no todo	05
25	Não atendimento ao tema proposto	10
	Total de pontos descontados:	-----

Resultado final:

100 menos os pontos descontados igual à média geral da redação _____

Ao utilizar uma série de critérios como estes, o professor desconta do escore máximo de 100 os pontos por penalidades. Há que se observar também que esta série de critérios apresenta pesos diferentes para os diferentes aspectos a serem julgados pelos avaliadores. Por exemplo, serão descontados 07 pontos no item *Ortografia* e 03 no item *Emprego de conectivos*, o que caracteriza uma típica série de critérios analíticos.

A maior desvantagem deste tipo de avaliação é o tempo exigido do avaliador. Este terá que fazer tantas leituras quantas forem necessárias para avaliar cada um dos itens que compõem o rol de critérios. Isto faz com que este tipo de avaliação se torne pouco prático, pois é necessário que os avaliadores disponham do tempo suficiente para a realização da tarefa. Todavia, segundo Hughes (1994 : 93), esta não é a única desvantagem da avaliação analítica. Há o sério risco de um avaliador se concentrar mais nos aspectos a serem considerados e perder de vista o texto como um todo. Conhecidos alguns critérios analíticos de avaliação e o que vem a ser a técnica analítica, passemos à técnica holística.

3.6.3. A Técnica Holística

Como já foi mencionado anteriormente, esta técnica consiste de uma leitura geral do texto, seguida da atribuição de uma nota, sem análise interpretativa das partes. Nela não se utilizam necessariamente critérios específicos e definidos para avaliar a expressão escrita. Avosani (1981 : 47), como conclusão de sua revisão bibliográfica sobre a correção de redações, afirma que "com base na sua experiência pessoal, no todo apresentado, os professores avaliadores atribuem notas ou um conceito sem a preocupação em detalhar ou analisar os componentes básicos da redação." O autor destaca ainda que tal técnica de avaliação peca pela subjetividade e generalização intuitiva dos julgamentos. Contudo, não demanda muito tempo de trabalho de cada avaliador. Este dá seu parecer após , no máximo, duas leituras. Tal agilidade faz com que um avaliador possa avaliar muito mais redações em um período de tempo relativamente curto e também facilita a organização de grupos de

avaliadores, garantindo assim a avaliação de um mesmo texto por mais pessoas, o que favorece uma avaliação mais objetiva.

Segundo Hughes (1994 : 97), uma avaliação holística bem organizada apresentará resultados de alta confiabilidade se houver quatro avaliadores para uma mesma redação, sendo a nota final a média das notas atribuídas por cada um. O referido autor menciona ainda que não há nada de mágico com relação ao número quatro. Segundo ele, pesquisas foram feitas e revelaram que redações julgadas por quatro avaliadores diferentes apresentam um maior grau de confiabilidade nos resultados. No entanto, embora teoricamente correto, este procedimento não é nada prático, pois na realidade dificilmente há a possibilidade de um grupo de quatro professores bem treinados trabalharem conjuntamente, avaliando as redações dos mesmos alunos. Convém mencionar ainda que, para uma avaliação holística apresentar resultados de alta confiabilidade, é preciso que se faça um treinamento dos professores avaliadores, com o intuito de evitar que diferentes pessoas avaliem diferentes aspectos de um mesmo texto, diminuindo assim a confiabilidade dos resultados.

Como exemplo de escala holística, vejamos o quadro 04.

Quadro 04 : Escala holística utilizada pela *Royal Society of Arts Examination in the Communicative Use of English as a Foreign Language*

FN	Padrão de falante nativo
FN-	Próximo ao padrão de falante nativo
MA	Claramente mais que adequado
MA-	Possivelmente mais que adequado
A	ADEQUADO PARA ESTUDAR NESTA UNIVERSIDADE
D	Duvidoso
I	Inadequado
DA	Distante da adequação

Hughes, 1994:87.

Esta escala é utilizada para avaliar a capacidade dos candidatos, em fase de admissão na *Royal Society of Arts*, de acompanhar os cursos, em inglês, daquela instituição. Tal escala se apresenta de maneira bastante subjetiva e foi criada baseando-se em textos produzidos por alunos de graduação daquela instituição e no julgamento do que seria aceitável, em termos de uso da língua, pelos seus professores. Para esta situação específica, a escala tem apresentado resultados satisfatórios; porém, é preciso fazer adaptações caso se deseje utilizá-la num contexto diferente.

Todavia, este tipo de escala puramente holística é difícil de ser encontrada; na maioria das vezes nos deparamos com sistemas de avaliação de escrita que combinam critérios holísticos com analíticos.

Antes de prosseguirmos, é interessante apresentarmos ainda alguns exemplos deste tipo de avaliação, desenvolvidos por instituições de renome internacional. A Universidade de Princeton , nos Estados Unidos, criou em 1986 o *Test of Written English (TWE)* como um apêndice do TOEFL (*Test Of English as a Foreign Language* - ver quadro 05). Estudantes do mundo inteiro que têm a intenção de ingressar em uma universidade norte-americana fazem o teste para verificar seu grau de proficiência em expressão escrita e para verificar se estão aptos para acompanharem cursos acadêmicos em inglês. A avaliação deste tipo de teste é cuidadosa e segue critérios bem específicos. Os avaliadores são submetidos a um treinamento antes do início do processo de avaliação. Não há nota mínima para ser aprovado no *TWE*, pois o mesmo tem a intenção de apontar apenas a qualidade da produção escrita dos alunos. Entretanto, como a maior parte das universidades norte-americanas exige que o aluno tenha um escore mínimo de 4 no *TWE* , muitas pessoas pensam ser esta a nota mínima para "ser aprovado".

Quadro 05: Critérios de avaliação no *TWE*

Escores	Explicação dos escores do <i>TWE</i>
6	Demonstra clara competência em ambos os níveis, retórico e sintático, ao escrever, embora o texto possa apresentar erros ocasionais.
5	Demonstra competência em ambos os níveis, retórico e sintático, ao escrever, embora o texto apresente erros ocasionais.
4	Demonstra uma competência mínima em ambos os níveis, retórico e sintático, ao escrever.
3	Demonstra uma competência em desenvolvimento ao escrever, mas o texto apresenta-se imperfeito, seja no nível retórico, seja no nível sintático, ou em ambos.
2	Sugere incompetência ao escrever.
1	Demonstra incompetência ao escrever.

Fonte: Hamp-Lyons (1989:75)

Os critérios acima foram criados para garantir que os textos produzidos pelos alunos oriundos de todas as partes do mundo possam ser avaliados com base em um padrão absoluto de redação acadêmica. O mesmo é aplicado para avaliação dos textos de todos os candidatos que se submetem a esta prova, independentemente do seu país de origem , de seu "background" ou de critérios locais de avaliação da expressão escrita. O escore reflete a qualidade do texto produzido.

É importante ressaltar que este tipo de critério de avaliação é o que Hamp-Lyons (1989 : 79) chamou de "avaliação holística modificada" , ou seja, os avaliadores que corrigem o *TWE* são treinados para julgarem os textos não apenas pela impressão geral que os mesmos lhes transmitem , mas também utilizando-se de critérios como competência retórica (*rheterical competence*) e competência sintática (*syntactic competence*). A competência retórica está relacionada à organização de idéias no texto e a competência sintática refere-se à utilização das regras gramaticais nas frases. Os avaliadores classificam os candidatos de acordo com uma série de habilidades para a produção de textos. Assim, um candidato que tenha recebido nota 4 no *TWE* demonstra uma competência mínima necessária para a comunicação escrita, com um mínimo necessário de competência sintática e retórica; ele está apto a acompanhar um curso de graduação e produzir redações acadêmicas sem grandes dificuldades. Julga-se a qualidade da produção escrita como um todo e a função comunicativa do texto é sempre privilegiada pela mescla de técnicas holísticas e analíticas.

Vejamos mais um exemplo de avaliação holística de textos numa escala produzida por Shohamy, Reves e Bejarano (Shohamy, 1985 : 59), da Universidade de Tel-Aviv ,em 1984:

Quadro 06: Escala para avaliação holística da Universidade de Tel -Aviv

0	Apenas uma tentativa.
1	Algum domínio de língua, texto quase incompreensível, com dificuldades de comunicação.
2	Texto compreensível, com alguns problemas gramaticais, morfológicos e/ou sintáticos. Há possibilidade de comunicação, mas com predominância de erros. Mensagem simplista, mas compreensível, sem indicação de sofisticação para expressão.
3	Texto comunicativo. Conhecimento de aspectos sócio-lingüísticos; boa sintaxe; linguagem idiomática; vocabulário adequado. Alguns erros gramaticais menores, porém sistemáticos.
4	Texto próximo ao de um falante-nativo.
5	Texto comparável ao de um falante nativo.

De acordo com esta escala de avaliação o aluno que obtiver o escore máximo é um escritor com proficiência comparável a de um falante nativo de língua inglesa. Para a avaliação de redações segundo esta escala, espera-se que os avaliadores leiam o texto não mais do que duas vezes para atribuírem suas notas, e isto com base na impressão geral que os textos despertam. Não é necessário fazer correções nos mesmos e somar partes para atribuir uma nota global. Os avaliadores lêem os

textos e os ordenam em relação uns com os outros de acordo com a posição que ocupam na escala de 0 a 5. Tal procedimento é adotado pelo fato de os níveis da escala serem muito abrangentes. Consequentemente, um texto é melhor, pior ou igual a outro e as notas são atribuídas com base na impressão despertada no leitor. Segundo Bachman (1991: 344-8), ao classificar os textos em níveis, estamos categorizando indivíduos de acordo com sua capacidade de desempenho das habilidades linguísticas, portanto, é necessário que se estabeleça o "ponto zero" e o "ponto perfeito" de cada um dos níveis para que possamos "nivelar" os textos a partir destes pontos de referência. Na verdade, estaremos relativizando, ou seja, comparando os textos a um padrão esperado para cada nível. Desta forma, as escalas até aqui apresentadas, em grande parte, são escalas relativas; exceção feita à escala do *TWE* e à da Universidade de Tel-Aviv, que são absolutas pois o escore obtido é aquele sob qualquer circunstância, como em um teste escrito de habilitação para motorista ou num teste de QI.

Por último, vejamos ainda um outro exemplo de escala de avaliação holística elaborado e adotado pela Universidade de Cambridge para o teste *CAE* (*Certificate of Advanced English*):

Quadro 07: Escala para avaliação da expressão escrita do CAE

5	Efeito total no leitor. Número mínimo de erros. Uso natural, controlado e talentoso da língua, demonstrando uma boa variedade de vocabulário e estruturas. Realização da tarefa: trabalho bem organizado, bom uso de conectivos, adequação de registro, nenhuma omissão relevante.
4	Texto suficientemente natural, ocorrendo erros apenas ao tentar usar uma linguagem mais elaborada. Algumas evidências de variedade de vocabulário e estrutura. Uma boa tentativa de realização da tarefa. As omissões, quando ocorrem, são menores. Atenção para a organização e coesão; registro nem sempre natural, mas com um efeito positivo sobre o leitor.
3	Uso satisfatório do inglês, embora com deficiências na variedade. Ocorrência de erros sérios ocasionais, que não impedem a comunicação, e que exigem paciência por parte do leitor. Tarefa razoavelmente cumprida, algumas tentativas de organização e coesão de frases que afetam negativamente o leitor.
2	Erros que às vezes dificultam a comunicação e/ou língua muito elementar. Tentativa de realização da tarefa, mas com omissões notáveis e/ou falta de organização e coesão que afetam negativamente o leitor.
1	Falta séria de controle e/ou erros sérios e frequentes. Pouca variedade de língua. Tentativa totalmente frustrada de realização da tarefa.
0	Linguagem insuficientemente compreensível para fins de avaliação.

Fonte: Aspinall & Hashemi, 1991.p. 34.

Weir (1993 : 164) afirma que o problema com este tipo de escala está no fato de que a mesma não foi desenvolvida a partir de experiências empíricas. Cada nível da escala parece representar diferentes níveis de proficiência, mas como ainda não se sabe exatamente a ordem de aquisição dos vários aspectos linguísticos avaliados num texto escrito, este nivelamento é feito

através dos textos produzidos pelos alunos e, com base nisso, são desenvolvidas as escalas de avaliação. Este tipo de técnica de avaliação não leva em conta que os candidatos possuem controle diferenciado dos diversos aspectos lingüísticos que envolvem a produção escrita, ou seja, um candidato pode ter maior domínio lexical do que sintático, ao passo que outro pode ter maior domínio sintático do que lexical, sendo que a ambos é atribuída a mesma nota. Na verdade, há desempenhos diferentes para os diferentes componentes da expressão escrita e isto não é considerado numa avaliação puramente holística. Nela não é possível detectar quais são os pontos fortes ou fracos dos candidatos, o que dificulta sua utilização na avaliação de redações.

3.7. Considerações Finais

Apresentadas as técnicas de avaliação de redação, ficamos com o dilema de optar por uma ou outra. Se a avaliação holística peca pela subjetividade e generalização intuitiva dos julgamentos, a avaliação analítica decompõe um todo e não é capaz de sintetizá-lo fidedignamente em termos de resultados. Não há uma única resposta e nem sempre existe a opção por uma ou outra técnica. Talvez melhor seria utilizar-se de ambas e trabalhar com o que Hamp-Lyons (1989 : 79) chamou de "holística modificada". Entendemos que a dicotomia "técnica holística/técnica analítica" não procede, pois mesmo utilizando-se de uma série de critérios analíticos de avaliação um avaliador ainda assim pode deixar-se levar por suas "impressões" e subjetividade na interpretação de tais critérios. Por outro lado, ao observarmos as escalas chamadas de holísticas pelos próprios autores das mesmas, constatamos que quase todas apresentam uma série de critérios específicos a serem considerados no momento da avaliação, característica esta típica da avaliação analítica.

Apesar de termos concentrado esforços em apresentar exemplos da técnica holística e da técnica analítica, nem sempre é necessário optar entre uma ou outra. Esta também parece ter sido a conclusão do *University of Cambridge Local Examination Syndicate* ao elaborar a nova escala de critérios utilizada para avaliação da expressão escrita do IELTS (*International English Language Testing Service*). Tal escala combina critérios analíticos com critérios holísticos e os resultados finais

são a média dos três elementos considerados pelos avaliadores: 1) Qualidade Comunicativa, 2) Argumentos, Idéias e Evidências e 3) Vocabulário e Estrutura da Frase . A avaliação se dá da seguinte forma: após uma primeira leitura do texto, o avaliador atribui uma nota para a Qualidade Comunicativa do mesmo (1), baseada na impressão global que o mesmo apresentou (ver Anexo 4). Em seguida, o mesmo avaliador lê novamente o texto para então atribuir notas para os itens 2 e 3, analiticamente. Como este é um teste que envolve não apenas a língua escrita, o resultado final é a média geral das notas atribuídas a cada uma das partes do teste (como acontece com todos os outros testes da instituição). Esta parece ser uma nova tendência, a de unir técnicas na tentativa de alcançar uma maior confiabilidade, sem relevar a praticidade e a validade da avaliação a um segundo plano. Entretanto, a nota é decidida por apenas um avaliador , o que vai contra a posição de Hughes (1994) de que ao menos dois avaliadores são necessários. Este avaliador único é escolhido pelo Conselho Britânico após um teste e tem um manual detalhado dos procedimentos a serem seguidos para a utilização da tabela.

Colocados os prós e os contras de ambas as técnicas, passamos ao capítulo seguinte, onde apresentaremos nossa pesquisa.

4. METODOLOGIA

4.1. Amostra

A fim de fazer a análise do problema exposto na parte introdutória do presente trabalho, foram avaliadas 20 redações, por 10 avaliadores, como segue: num primeiro momento utilizou-se a técnica holística e, posteriormente, a técnica analítica.

As redações foram elaboradas pelos alunos da disciplina de Língua Inglesa Escrita III, do curso de graduação em Letras da Universidade Federal do Paraná (UFPR), do primeiro semestre do ano letivo de 1996. A escolha desta disciplina deu-se com base no seu programa, que estabelece que no final do semestre letivo, os alunos devem ser capazes de redigir um texto coerente, "ensaios", na língua estrangeira, e não mais apenas parágrafos isolados. Dos trinta alunos desta disciplina, vinte foram selecionados aleatoriamente.

Os avaliadores, por sua vez, foram escolhidos de acordo com sua experiência profissional, ou seja, deviam ser pessoas com licenciatura plena em Letras-Inglês com, no mínimo, três anos de experiência no ensino de línguas e que estivessem familiarizados com a estrutura curricular do curso de graduação de Letras da UFPR. De um total de quinze professores dispostos a tomar parte na pesquisa, que atendiam a todos estes requisitos e residiam na região sul do país, dez foram escolhidos aleatoriamente.

4.2. Procedimentos para a coleta de dados

4.2.1. Com os alunos

Com o intuito de preparar os alunos para a elaboração das redações houve, primeiramente, uma conversa informal entre os professores e alunos integrantes das duas turmas da disciplina de

Escrita III. Em seguida lhes foi proposto, pelos respectivos professores, o tema a ser elaborado : "O Vestibular". Como etapa seguinte foram discutidas as idéias e opiniões por eles apresentadas que, na seqüência, foram ordenadas num esboço ("outline") sobre o tema. Somente então foi solicitado aos alunos que elaborassem suas redações individuais, cuja extensão não deveria ultrapassar vinte (20) linhas. O tempo disponível para a realização da tarefa foi de quarenta (40) minutos. Na etapa seguinte, as vinte redações foram datilografadas, para evitar que a caligrafia dos alunos interferisse, de alguma forma, no processo de avaliação (ver anexo 1).

4.2.2. Com os professores

Cada um dos dez professores avaliadores recebeu dois envelopes, cada um contendo as mesmas vinte redações datilografadas e orientações sobre como proceder em cada uma das duas fases do experimento (ver Anexo 2). Nenhum outro tipo de treinamento ou informação foi dada aos professores antes, durante ou após a realização da tarefa para evitar a nossa interferência no processo. Na primeira fase os professores deveriam avaliar as redações do primeiro envelope de acordo com a técnica holística e, para tanto, seguir as instruções dadas. Deveriam ler cuidadosamente cada uma das redações, no máximo duas vezes, sem se ater a detalhes tais como repetição de palavras, orações ou períodos, julgando o texto de acordo com sua impressão global e atribuindo-lhe uma nota de zero a cem.

Nesta fase não foram fornecidos critérios para a atribuição de notas e o professor deveria utilizar apenas a impressão geral que cada uma das redações despertou, baseando-se em sua própria hierarquia de valores. As únicas informações que os professores receberam até esse momento foram:

- Nível dos alunos autores das redações : Escrita III
- Tipo de escrita: acadêmica
- Função do texto: argumentativo
- Tempo para a realização da tarefa: 40 minutos

- Comprimento do texto: 20 linhas
- Propósito da atividade: ajudar a organizar o material para o curso, ajudar a diagnosticar os pontos fortes e fracos dos alunos de modo a possibilitar um acompanhamento no desenrolar do curso.

Ao término desta fase, os professores deveriam fazer um intervalo de, no mínimo, vinte e quatro horas, para iniciar a fase seguinte. Com isto procurou-se evitar que fatores como o cansaço físico ou mental e a interferência de uma técnica sobre a outra influenciassem o processo, já que estas variáveis não seriam passíveis de controle.

No início da segunda fase os professores não sabiam que iriam encontrar as mesmas redações da primeira fase. Houve ainda a preocupação de dispor as redações de maneira diferente daquela encontrada no primeiro envelope, mas sem nenhum padrão definido.

Diferentemente da primeira fase, neste segundo momento os professores foram devidamente orientados de como proceder para a avaliação analítica, cujos critérios lhes foram fornecidos. Primeiramente eles deviam familiarizar-se com o processo e só então ler as redações, observando a idéia geral e a mensagem transmitida por cada uma. Feito isto, deviam procurar os itens "Conteúdo" e "Organização" nas tabelas de critérios e tentar atribuir um escore para cada um deles, com base na qualidade apresentada pelo texto e com a escala estabelecida previamente. A seguir, o professor devia reler a redação e atribuir escores aos itens "Vocabulário", "Uso da Língua" e "Mecânica da Língua". A nota final da redação seria a soma das cinco notas atribuídas a cada um dos itens individualmente (ver Anexo 3).

Como é possível perceber, o professor teve que fazer três leituras de cada uma das vinte redações, antes de atribuir-lhes uma nota final. Além disso, cada um dos cinco itens possuía os detalhes aos quais o professor devia ater-se ao julgar um componente específico. Esses detalhes, que eram reconhecidos por especialistas em comunicação escrita (Jacobs et al. , 1981), foram sintetizados em uma tabela (ver Anexo 2). A opção por esta tabela de critérios, entre outras, deu-se pelo fato de a mesma ser bastante completa e de fácil utilização. Por exemplo, ao avaliar o item "Mecânica da Língua" o professor devia estar atento para as convenções do idioma, os erros de ortografia, pontuação e capitalização. Devia estar atento também para a utilização de parágrafos e o aluno devia

ser classificado em categorias, recebendo a nota referente ao seu desempenho. Para um desempenho de excelente a muito bom a nota seria 5, de bom a mediano a nota seria 4, de razoável a fraco, 3 e muito fraco, 2.

Como os critérios foram previamente determinados e estavam claros, não surgiu a necessidade de nenhum esclarecimento durante o processo. O prazo total dado aos professores para a avaliação das redações foi de dez dias.

Antes de prosseguirmos, é importante ressaltar que os professores não foram treinados para utilizar a tabela de critérios que deveriam aplicar. Embora este procedimento seja contrário aos ensinamentos e sugestões encontrados na revisão da literatura, o não treinamento dos professores foi proposital, pois o objetivo era verificar os resultados obtidos através da utilização das diferentes técnicas e, além disso, esta é a realidade na UFPR.

4.3. Análise dos dados

Os dados foram constituídos pelas notas atribuídas pelos professores às 20 redações, segundo as técnicas holística e analítica. A análise dos dados foi realizada em várias etapas. Primeiramente as notas foram condensadas em tabelas para cada professor, a partir das quais se calculou o coeficiente de correlação para verificar se as notas atribuídas segundo uma técnica acompanhavam as notas atribuídas segundo a outra. Este procedimento foi realizado para os dez professores, separadamente e em conjunto. Finalmente, foi feita uma análise de variância para verificar os fatores que influenciaram a atribuição das notas. A análise e a interpretação dos dados estão no próximo capítulo do nosso trabalho.

5. ANÁLISE E INTERPRETAÇÃO DOS DADOS

5.1. Visão Geral

De acordo com o exposto no item 4.3., neste capítulo faremos primeiramente a apresentação das notas atribuídas pelos professores às 20 redações segundo as técnicas de avaliação holística e analítica. Em seguida, faremos a transformação das notas atribuídas segundo a técnica analítica para serem comparáveis às da técnica holística. A partir daí, então, serão computados os coeficientes de correlação entre as notas das duas técnicas e, finalmente, os dados serão submetidos a uma análise de variância para verificar a ação isolada ou em conjunto das variáveis independentes (professor, redação e técnica de avaliação) sobre a dependente (notas atribuídas).

5.2. Apresentação dos Dados

Para a análise dos dados coletados, as notas atribuídas pelos professores às vinte redações, segundo a técnica holística e analítica, foram previamente ordenadas em tabelas. A Tabela 1 apresenta as notas segundo a técnica holística e a Tabela 2 segundo a técnica analítica.

Tabela 1: Notas atribuídas pelos 10 professores às 20 redações pela Técnica Holística

Redação	Professores									
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
R1	30	50	30	62	70	48	70	45	50	72
R2	40	50	25	78	65	50	50	45	40	68
R3	40	60	46	80	68	57	70	60	70	60
R4	70	70	60	70	75	60	75	80	60	83
R5	60	70	48	80	75	70	70	70	80	67
R6	40	45	55	60	75	54	65	60	50	80
R7	50	70	35	48	70	70	70	65	60	60
R8	80	75	25	82	70	68	75	65	80	65
R9	60	80	58	92	70	74	75	70	70	63
R10	30	50	32	73	75	45	60	55	50	80
R11	40	60	37	63	65	55	60	50	70	50
R12	40	85	40	57	70	79	75	75	90	75
R13	40	45	33	47	60	40	70	50	50	50
R14	45	55	60	50	68	52	70	70	60	53
R15	10	40	31	48	60	30	50	45	40	40
R16	50	70	42	40	65	65	75	60	60	55
R17	30	70	39	53	70	50	70	60	40	77
R18	40	80	39	85	78	75	80	55	80	88
R19	90	90	80	93	85	85	85	95	100	93
R20	20	50	40	45	65	40	60	55	50	73

Tabela 2: Notas atribuídas pelos 10 professores às 20 redações pela Técnica Analítica

Redação	Professores									
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
R1	42	65	53	41	60	53	68	51	64	74
R2	42	68	53	67	75	55	57	56	51	67
R3	41	71	51	63	81	60	77	68	61	73
R4	73	80	69	82	90	60	78	70	100	86
R5	55	75	65	65	88	74	75	65	85	70
R6	45	74	64	53	75	60	70	67	70	83
R7	81	79	64	52	72	78	87	66	86	61
R8	85	77	62	76	90	74	91	72	91	76
R9	76	83	61	86	81	78	90	78	90	62
R10	61	83	53	81	81	52	88	55	96	84
R11	68	60	55	69	78	64	71	62	82	68
R12	58	83	75	81	88	85	84	68	87	70
R13	55	70	62	56	74	55	71	56	71	66
R14	72	80	70	48	72	58	62	57	57	63
R15	38	61	61	42	73	43	60	58	51	47
R16	72	84	65	47	77	72	89	56	75	62
R17	59	72	64	44	88	60	67	63	75	72
R18	56	84	62	74	92	84	92	69	97	81
R19	95	98	91	91	98	94	96	94	98	90
R20	50	73	62	63	64	56	70	52	62	72

Como é possível observar através da metodologia de ambas as técnicas de correção das redações, na técnica holística a nota da redação podia variar de zero a cem. Porém, para a técnica analítica, a nota só podia variar entre trinta e quatro e cem, o que pode causar uma falsa impressão de que as notas dadas pela segunda técnica são maiores, quando na realidade é a diferença entre a nota máxima e mínima nesta técnica que é menor. Para que os dados obtidos segundo a primeira técnica fossem comparáveis às da segunda, fez-se uma transformação dos dados da técnica analítica para valores percentuais, segundo a fórmula:

$$Z = ([X - 34] \cdot 100) : 66$$

Segundo esta fórmula, de cada valor (X) foi subtraído o valor 34, que é o menor valor possível segundo a tabela de critérios de Jacobs et al. (1981). Consequentemente, a variação passa a ser de zero a sessenta e seis. O valor resultante da subtração foi dividido por 100, o que resultou na porcentagem. Os resultados da transformação das notas segundo a técnica analítica estão contidos na Tabela 3.

Tabela 3: Notas segunda a Técnica Analítica transformadas para valores percentuais

Redação	Professores									
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
R1	12,1	47	28,8	10,6	39,4	28,8	51,5	25,8	45,5	60,6
R2	12,1	51,5	28,8	50	62,1	31,8	34,9	33,3	25,8	50
R3	10,6	56,1	25,8	43,9	71,2	39,4	65,2	51,5	40,9	59,1
R4	59,1	69,7	53	72,7	84,9	39,4	66,7	54,6	100	78,8
R5	31,8	62,1	47	47	81,8	60,6	62,1	47	77,3	54,6
R6	16,7	60,6	45,5	28,8	62,1	39,4	54,6	50	54,6	74,2
R7	71,2	68,2	45,5	27,3	57,6	66,7	80,3	48,5	78,8	40,9
R8	77,3	65,2	42,4	63,6	84,9	60,6	86,4	57,6	86,4	63,6
R9	63,6	74,2	40,9	78,8	71,2	66,7	84,9	66,7	84,9	42,4
R10	40,9	74,2	28,8	71,2	71,2	27,3	81,8	31,8	93,9	75,8
R11	51,5	39,4	31,8	53	66,7	45,5	56,1	42,4	72,7	51,5
R12	36,4	74,2	62,1	71,2	81,8	77,3	75,8	51,5	80,3	54,6
R13	31,8	54,6	42,4	33,3	60,6	31,8	56,1	33,3	56,1	48,5
R14	57,6	69,7	54,6	21,2	57,6	36,4	42,4	34,9	34,9	43,9
R15	6,1	40,9	40,9	12,1	59,1	13,6	39,4	36,4	25,8	19,7
R16	57,6	75,8	47	19,7	65,2	57,6	83,3	33,3	62,1	42,4
R17	37,9	57,6	45,5	15,2	81,8	39,4	50	43,9	62,1	57,6
R18	33,3	75,8	42,4	60,6	87,9	75,8	87,9	53	95,5	71,2
R19	92,4	97	86,4	86,4	97	90,9	93,9	90,9	97	84,9
R20	24,2	59,1	42,4	43,9	45,5	33,3	54,6	27,3	42,4	57,6

Um primeiro exame das duas tabelas revela que, quanto à técnica holística (Tabela 1), há variabilidade entre os professores na avaliação de uma redação. Tomemos como exemplo as notas da redação 1 que variam de 30 a 72, sendo que o escore 30 foi atribuído por dois professores, o mesmo ocorrendo com os escores 50 e 70. Já na redação 19 as notas se concentram num intervalo bem menor, isto é, entre 80 e 100, tendo o escore 90 frequência 2 (atribuído por dois professores) e o escore 85 a frequência 3 (atribuído por três professores). Na tabela 3 (técnica analítica convertida para valores percentuais) também se verifica uma variabilidade de professor para professor na avaliação de uma redação. Quanto à redação 1, os escores variam de 10,6 a 60,6, sendo que nenhum escore tem frequência superior a 1; quanto à redação 19, os escores também apresentam uma variabilidade bem menor, variando estes entre 84,9 e 97, sendo que o último foi atribuído por três professores. Comparações semelhantes podem ser feitas com relação às demais redações segundo cada uma das técnicas de avaliação.

5.3. A Correlação

Sendo que um dos objetivos do presente trabalho foi comparar as duas notas que o mesmo professor atribuiu a uma redação segundo a técnica holística e a analítica, optou-se por calcular uma medida associativa entre estas notas. Para verificar se as notas atribuídas por cada professor segundo a técnica analítica acompanhavam as da técnica holística, decidiu-se calcular o Coeficiente de Correlação por Postos de Spearman (r_s), que exige uma escala de mensuração, no mínimo, ordinal. Verificou-se, em seguida, se o coeficiente era significativo ou não. Sempre que possível, as tabelas contendo os resultados, segundo as duas técnicas, foram complementadas por gráficos que apresentam a mesma informação, desta vez na forma de diagramas de dispersão.

5.3.1. Coeficiente de correlação entre os escores da técnica holística e analítica para cada professor

Para o cálculo dos coeficientes de correlação primeiramente foram atribuídos postos às redações, com base nas notas atribuídas, segundo cada técnica. O resultado foi a ordenação das 20 redações, para cada técnica, de sorte que à redação com a melhor nota fosse atribuído o posto 20 e à redação com a nota mais baixa o posto 1. Quando duas ou mais redações obtiveram a mesma nota, ou seja, quando houve empate de notas, foi-lhes atribuída a média dos postos que lhes caberiam se não tivesse havido empate. Desta forma, se a nota mais baixa para um determinado professor foi 45 e duas redações tiveram esta nota, ambas receberam o escore 1,5 que corresponde à média entre os postos 1 e 2.

A Tabela 4 apresenta as 20 notas atribuídas pelo Professor 1, segundo a técnica holística, em ordem crescente, juntamente com as 20 notas segundo a técnica analítica.

Tabela 4: Notas e postos para as 20 redações do Professor 1

Redação	Nota Holística	Posto Holístico	Nota Analítica	Posto Analítico
15	10	1	38	1
20	20	2	50	6
1	30	4	42	3,5
10	30	4	59	11
17	30	4	61	12
2	40	9	41	2
3	40	9	42	3,5
6	40	9	45	5
11	40	9	55	7,5
12	40	9	56	9
13	40	9	58	10
18	40	9	68	13
14	45	13	72	14,5
7	50	14,5	72	14,5
16	50	14,5	81	18
5	60	16,5	55	7,5
9	60	16,5	76	17
4	70	18	73	16
8	80	19	85	19
19	90	20	95	20

Tabelas semelhantes foram elaboradas para os demais professores. A correlação entre a nota dada pela técnica holística por um determinado professor e a nota dada pela técnica analítica

pelo mesmo professor será perfeita se os postos atribuídos às redações pelo primeiro método forem iguais aos do segundo método; ou seja, se $X_i = Y_i$, para todos os i 's. Isto é o mesmo que dizer que $X_i - Y_i = 0$.

Chamando as diferenças entre os postos, segundo as duas técnicas, d_i , é possível utilizá-las como medida de disparidade entre os dois conjuntos de postos. A magnitude dos diversos d 's dá uma idéia do relacionamento entre as duas notas da mesma redação. Se a relação entre os dois conjuntos de postos fosse perfeita, todos os d 's teriam valor zero. Assim, quanto maiores os diversos d 's, menos perfeita é a associação entre as duas técnicas de avaliação.

Para o cálculo do coeficiente r_s utilizou-se d^2 , neutralizando o efeito dos d 's negativos, pois a relação de que, quanto menor o "d", mais perfeita é a relação entre os conjuntos é mantida quando se utiliza seu valor ao quadrado. A fórmula para o Coeficiente de Correlação por Postos de Spearman com observações empatadas é:

$$r_s = \frac{\sum x^2 + \sum y^2 - \sum d^2}{2\sqrt{\sum x^2 \sum y^2}}, \text{ onde}$$

$$\sum x^2 = \frac{N^3 - N}{12} - \sum T_x \quad \text{e} \quad \sum y^2 = \frac{N^3 - N}{12} - \sum T_y$$

Nesta fórmula:

N é o tamanho da amostra,

T é o fator de correção para postos empatados e é $T = t^3 - t$
12,

t é o número de observações empatadas em determinado posto,

$\sum d^2$ é a soma das diferenças para as 20 notas de redações de um professor.

Os resultados para o Professor 1 estão na tabela 5.

Tabela 5: Cálculo do "d" e "d²" para as 20 redações do Professor 1

Redação	Posto - Técnica Holística	Posto - Técnica Analítica	d	d ²
15	1	1	0	0
20	2	6	-4	16
1	4	3,5	0,5	0,3
10	4	11	-7	49
17	4	12	-8	64
2	9	2	7	49
3	9	3,5	5,5	30,3
6	9	5	4	16
11	9	7,5	1,5	2,3
12	9	9	0	0
13	9	10	-1	1
14	9	13	-4	16
18	13	14,5	-1,5	2,3
7	14,5	14,5	0	0
16	14,5	18	-3,5	12,3
5	16,5	7,5	9	81
9	16,5	17	-0,5	0,3
4	18	16	2	4
8	19	19	0	0
19	20	20	0	0

Para aplicar a fórmula do Coeficiente de Spearman tem-se primeiro que calcular $\sum x^2$ e $\sum y^2$:

$$\sum x^2 = \frac{20^3 - 20}{12} - \left(\frac{3^3 - 3}{12} + \frac{7^3 - 7}{12} + \frac{2^3 - 2}{12} + \frac{2^3 - 2}{12} \right) = 665 - 31 = 634$$

$$\sum y^2 = \frac{20^3 - 20}{12} - \left(\frac{2^3 - 2}{12} + \frac{2^3 - 2}{12} + \frac{2^3 - 2}{12} \right) = 665 - 1,5 = 663,5$$

Na expressão acima, cada termo entre parênteses é um dos conjuntos de empate e o número que está sendo elevado ao cubo é o número de elementos empatados. Sabendo-se que $d^2 = 343,8$ tem-se então:

$$r_s = \frac{634 + 663,5 - 343,8}{2\sqrt{634 * 663,5}} = \frac{953,7}{1297,16} = 0,735$$

que é o coeficiente de correlação por postos de Spearman para o primeiro professor.

Calculando-se o coeficiente entre as duas notas para todos os outros professores, como no exemplo anterior, para o Professor 1, obtemos os resultados que seguem na Tabela 6, a seguir:

Tabela 6: Resultados dos Coeficientes de Correlação por Postos de Spearman para os dez professores

Técnica Holística/Analítica, professor	Amostra N (número de redações)	Coeficiente r_s	Teste Estatístico t (N - 2)	Nível de Significância p - level
HOLP1 & ANAP1	20	.735450	4.605001	.000220
HOLP2 & ANAP2	20	.692521	4.072808	.000714
HOLP3 & ANAP3	20	.617993	3.335001	.003685
HOLP4 & ANAP4	20	.694654	4.097016	.000677
HOLP5 & ANAP5	20	.634359	3.481532	.002664
HOLP6 & ANAP6	20	.959538	14.45763	.000000
HOLP7 & ANAP7	20	.761708	4.987739	.000095
HOLP8 & ANAP8	20	.699752	4.155735	.000594
HOLP9 & ANA9	20	.617427	3.330062	.003726
HOLP10 & ANAP10	20	.825170	6.197540	.000008

O coeficiente de correlação (Tabela 6) teoricamente pode variar entre 0 e 1. Quanto mais próximo de 1, mais perfeita é a relação entre as duas notas. É possível observar, na tabela acima, que todos os coeficientes estão acima de 0,60 podendo-se então dizer que as notas atribuídas segundo ambas as técnicas se comportam de forma similar. Porém, apesar de todos os coeficientes apresentarem valores altos, alguns são altíssimos, como é o caso dos professores 6 ($r = 0,96$) e 10 ($r = 0,82$). Em oposição a isto, os resultados relativos a outros professores, como o número 3 ($r = 0,62$) e o número 9 ($r = 0,62$), por exemplo, apresentam coeficientes menores. Isto indica haver uma diferença de professor para professor quanto a concordância entre suas correções, de acordo com a técnica utilizada.

Após termos feito a ordenação por postos ("ranqueamento") das vinte redações, pelos dez professores, podemos afirmar que a redação 15 é a "pior" e a 19 é a "melhor" de todas. Por mais diferentes que sejam os posicionamentos dos postos intermediários entre estas duas redações, para cada um dos professores, todos concordaram que estes seriam os dois extremos do conjunto avaliado, independentemente da técnica utilizada para a avaliação. Segundo Alderson et al. (1995: 111), estas seriam exemplos de redação de consenso (*consensus script*), a de número 19, e a redação-problema (*problem script*), a de número 15. Não é objetivo do nosso trabalho definir o que seja um bom texto,

mas, com certeza, seria interessante estudar a fundo estas duas redações para tentar levantar características das mesmas, com respeito às quais os dez professores foram unânimes, embora trabalhando individualmente e sem treinamento prévio.

Cumpre, pois, ressaltar que a ordenação das redações pelos dez professores não foi homogênea, exceção feita às redações 15 e 19. Esta falta de homogeneidade não foi constatada apenas entre professores diferentes, mas até por parte de um mesmo professor, ao utilizar-se de técnicas diferentes de avaliação. Por exemplo, a redação número 17, segundo o professor 1, recebeu nota 30 e ocupou o 40 posto na avaliação holística. Porém, esta mesma redação, corrigida pelo mesmo professor utilizando-se da técnica analítica, recebeu nota 61 e ocupou o 12o lugar. Isto significa que o aluno, autor desta redação, poderia ter sido classificado ou não em um teste de qualificação, dependendo de quem fosse seu avaliador e da técnica utilizada. Além do mais, se observarmos apenas as notas, ficaremos estarecidos ao verificar que um mesmo professor é capaz de atribuir notas tão dispares para uma mesma redação, 30 segundo a impressão global e 61, seguindo critérios específicos (ver Tabela 5). Este fato já foi observado por outros autores, como, por exemplo, Hughes (1994), Hamp-Lyons (1990), Weir (1993) e Alderson et al. (1995).

Se de um lado houve discrepâncias acentuadas com relação a determinadas redações, em termos gerais, de outro lado as notas atribuídas por cada um dos professores segundo uma técnica de avaliação acompanhavam as da outra (ver Tabela 6). Para verificar se o coeficiente de correlação para cada professor era significativo, foi aplicado um teste estatístico de significância para a verificação dos resultados.

O teste utilizado foi o teste "**t**" de Student, onde a hipótese a ser testada era:

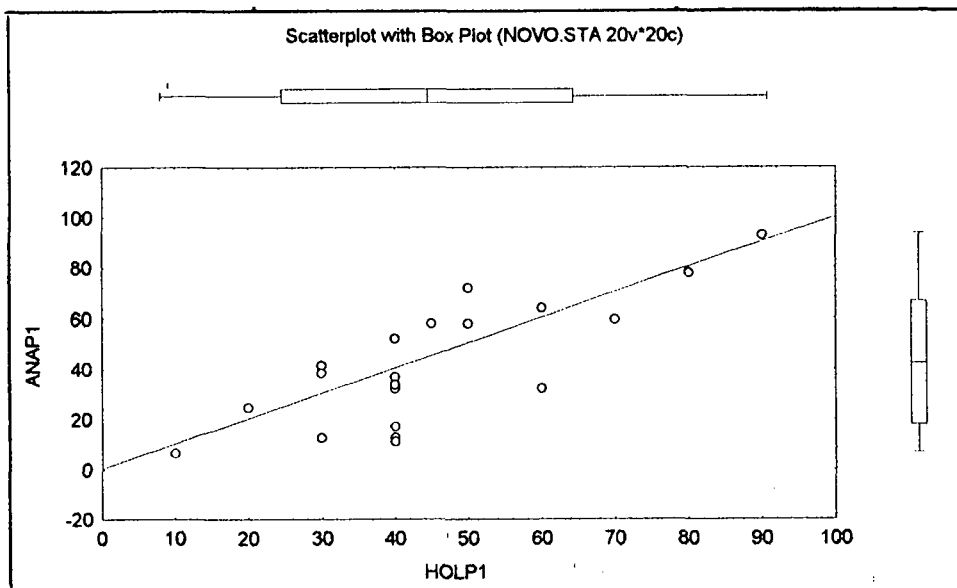
H₀ = Não existe correlação entre as duas técnicas para o professor em questão; e

H₁ = Existe correlação entre as duas técnicas para o professor em questão.

Com base nos resultados apresentados na Tabela 6, pode-se admitir que a um nível de 1% de significância, que é maior a todos os valores "p" encontrados, a hipótese nula (**H₀**) é rejeitada; ou seja, os dois métodos de avaliação podem ser considerados concordantes em termos estatísticos. Os

gráficos de número 1, a seguir, e de 2 a 10 em anexo (ver anexo 5), são diagramas de dispersão que apresentam a informação da Tabela 6, visualmente, para cada professor.

Gráfico 1 - Notas pelas técnicas holística e analítica para o Professor 1



Ao observarmos o valor dos coeficientes (Tabela 6), o Gráfico 1 e o Anexo 5, observa-se claramente que existe diferença na intensidade entre as correlações, conforme o professor, e que em alguns casos há uma tendência a dar maior nota segundo determinada técnica e menor por outra. Tal é o caso dos Professores 4, 6, 8 e 10 (Anexo 5) que, em média, deram notas maiores segundo a técnica holística; em oposição a isto, o Professor 9 superestimou a nota pela técnica analítica (Anexo 5). Segundo Weir (1993) este tipo de resultado é interessante porque nos mostra como a avaliação de uma redação por um professor não bem treinado pode ser frágil.

A seguir, foi calculado o coeficiente de correlação entre os escores das duas técnicas para todos os professores juntos.

5.3.2. Coeficiente de Correlação entre os escores da Técnica Analítica e Holística para todos os professores

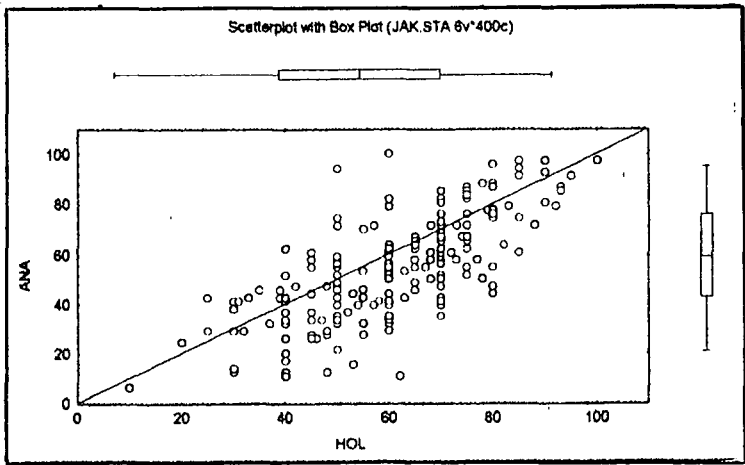
Neste caso, a metodologia segue os mesmos passos da utilizada no item anterior, com a diferença de que os postos são considerados para todos os professores juntos, com o intuito de verificar a relação entre os escores referentes às duas técnicas de avaliação. Os resultados estão na Tabela 7.

Tabela 7: Resultado do Coeficiente de Correlação por Postos de Spearman, para todos os professores juntos

	Amostra N	Coeficiente r_s	Teste Estatístico $t(N - 2)$	Nível de Significância p - level
HOL & ANA	200	.738498	15.41193	0.00

Na sequência, vejamos o Gráfico 2, que é o diagrama de dispersão dos dados correspondente à Tabela 7.

Gráfico 2 - Notas pela Técnica Holística e Analítica para todos os Professores



Segundo a análise dos dados referentes a todos os professores em conjunto chegamos à mesma conclusão a que chegamos para os professores separadamente, ou seja, de que,

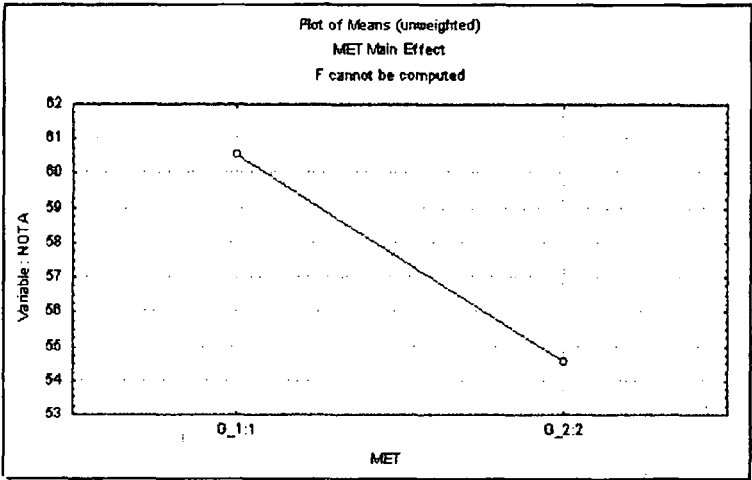
estatisticamente, as duas técnicas de avaliação apresentam resultados concordantes. Entretanto, pode-se aqui levantar a hipótese de que o que ocorre com os Professores 4, 6, 8 e 10 (atribuem maiores notas segundo a técnica holística) prevalece de uma maneira geral. Para nos certificarmos disto é necessário realizar uma Análise de Variância. É através desta análise que poderemos dizer se o professor, a redação e a técnica influenciam isoladamente ou de modo interativo a nota final. Como etapa preliminar disto, foram feitas análises com as médias aritméticas: por cada professor, por técnica e por redação.

5.4. As Médias

Para a Técnica Holística as médias encontram-se no Anexo 6 e para a Técnica Analítica no Anexo 7.

O Gráfico 3, abaixo, traz as médias da nota geral de cada técnica, obtida através da soma das notas atribuídas pelos dez professores às vinte redações, segundo cada uma das técnicas, e dividindo-se o resultado obtido por 20. Desta forma verifica-se o efeito da técnica de correção sobre a nota, independentemente do efeito do fator "Professor".

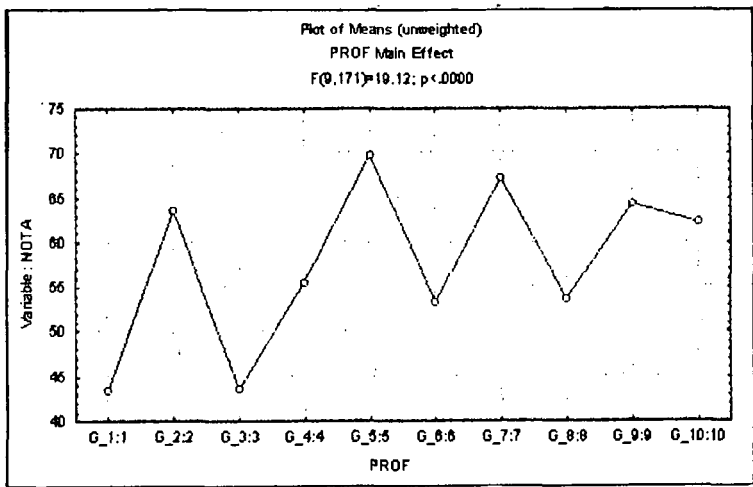
Gráfico 3 - Média da Nota Geral pelas Técnicas Holística e Analítica



Ao observarmos este gráfico podemos verificar que , embora a linha que une as duas médias apresente declive bastante acentuado, numericamente as duas médias não são muito distantes entre si, pois a média para as notas segundo a técnica holística é 60,52 (G1:1) e para a analítica é 54,36 (G2:2). Se esta diferença entre médias é significativa ou não só poderemos verificar com a Análise de Variância (ver item 5.5.).

O Gráfico 4 mostra a nota geral que foi obtida somando-se as médias segundo a técnica holística com as médias segundo a técnica analítica, por professor, e dividindo-se o resultado por dois. Aqui verifica-se o efeito do fator Professor sobre a nota, independentemente da técnica utilizada.

Gráfico 4 - Média da Nota Geral dada por cada Professor



Os mesmos resultados também são apresentados pela Tabela 8, a seguir.

Tabela 8: Média da nota geral por cada professor

Professor	Média Geral
1	43,23
2	63,43
3	43,43
4	55,40
5	69,73
6	53,23
7	67,08
8	53,60
9	64,20
10	61,10

Analisando os resultados do Gráfico 4 e da Tabela 8, notamos que o Professor 1 e o Professor 3 atribuíram notas bem mais baixas às 20 redações que seus colegas. Por outro lado, os Professores 5 e 7 parecem ter sido mais benevolentes em seus julgamentos, pois suas notas ficaram acima da faixa intermediária, formada pelos Professores 2, 4, 6, 8, 9 e 10. Podemos, então, afirmar que 60% dos professores atribuíram notas que convergiram em médias distribuídas de forma homogênea em torno da média geral de 57,44 $[(60,52 + 54,36) : 2 = 57,44]$, sendo 60,52 a média geral obtida segundo a técnica holística e 54,36 a média geral obtida segundo a técnica analítica (ver Gráfico 3). O restante dos professores, 40%, ou fica acima da média do grupo ou abaixo da mesma. No nosso entender este fato ocorreu devido à inadequação das nossas instruções para os professores. Embora a maioria, 60% dos professores, tenham considerado as instruções claras, 40% dos mesmos não se sentiram seguros ao utilizarem a escala de critérios adotada na pesquisa. Talvez se estes professores tivessem recebido um treinamento prévio, os resultados fossem mais homogêneos e a confiabilidade dos mesmos seria maior. Isto equivale a dizer que o não-treinamento dos professores afeta, sim, a confiabilidade dos resultados. Esta situação confirma a necessidade de treinamento tão recomendada por autores como Alderson et al. (1995) que frisam que o treinamento contínuo de professores é uma das maneiras de minimizar os desvios, para cima ou para baixo, da tendência geral.

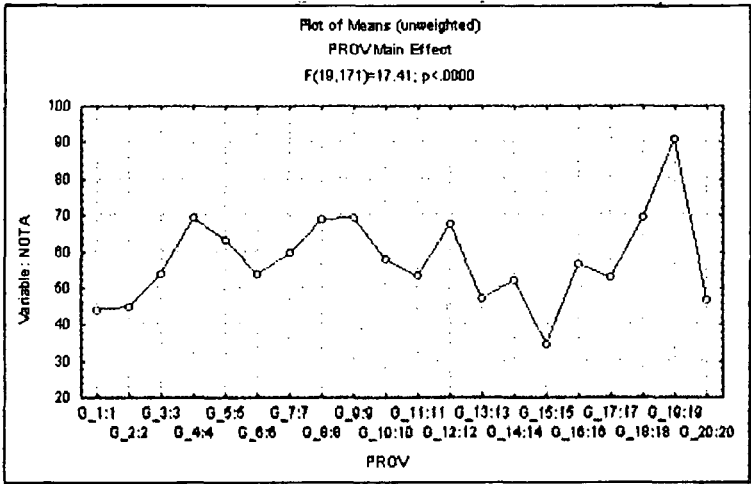
Vejamos agora o que nos mostram a tabela e o gráfico referentes às médias gerais de cada uma das redações.

Tabela 9 : Médias Gerais por Redação

Redação	Média Geral
1	43,86
2	44,57
3	53,74
4	69,10
5	63,07
6	53,53
7	57,15
8	68,65
9	69,32
10	57,35
11	53,03
12	67,56
13	46,68
14	51,81
15	34,40
16	56,30
17	52,50
18	69,17
19	90,64
20	46,42

O Gráfico 5 apresenta as mesmas informações em forma de diagrama de dispersão:

Gráfico 5 - Nota Geral atribuída a cada redação



O gráfico 5 e a Tabela 9 nos mostram as médias gerais (holística mais analítica dividida por dois), obtidas por cada uma das vinte redações. Aqui verificou-se o efeito do fator

"Redação" sobre a nota, independentemente dos fatores "Professor" e "Técnica de Correção". A análise dos mesmos confirma a informação já expressa no item 5.3.1. de que a Redação 19 é realmente a "melhor", pois obteve a maior média geral, e a Redação 15 é a "pior", pois obteve a menor média geral. Entretanto, esta interpretação simplista não é suficiente. É importante verificar se houve alguma interação entre os fatores de sorte que, alterando-se um deles, altera-se também a influência de outro sobre a variável dependente (nota). Por exemplo, o efeito da técnica de avaliação sobre a nota difere de acordo com o professor que corrige a redação? Para responder a esta e outras perguntas deste gênero, é necessário examinar a interação dos fatores já citados, através da Análise de Variância.

5.5. Análise de Variância

Como já foi frisado no item 5.3.2., a Análise de Variância nos possibilita verificar a relação das variáveis independentes (técnica, professor e redação) com a dependente (notas), esclarecendo assim como agem as variáveis independentes, em conjunto ou separadamente, sobre a variável dependente.

Este tipo de análise tem por objetivo responder a perguntas como "Será a variação entre as notas dadas pelos diferentes professores nas diferentes técnicas de avaliação apenas obra de variações aleatórias? Ou pode uma redação, dependendo do professor e da técnica que ele utiliza, receber uma nota maior ou menor?". Para responder a perguntas deste tipo é necessário, primeiramente, definir o delineamento do trabalho, ou seja, como o mesmo foi estruturado e que testes foram aplicados a partir deste delineamento.

Na presente pesquisa a análise de variância permitiu um delineamento do tipo **Split Plot**. Esta é a forma da análise quando se tem um experimento multifatorial, mas com restrições quanto a aleatorização de um dos fatores. No nosso trabalho os fatores foram o Professor, a Redação e a Técnica de correção, sendo que não houve aleatorização quanto ao último fator, isto é, quanto à

seqüência das técnicas de correção. Primeiramente todos os professores corrigiram todas as redações pela técnica holística, e depois todos os professores corrigiram todas as redações pela técnica analítica. Dentro de cada técnica a seqüência de correção das redações foi aleatorizada para cada professor.

É importante definir o tipo de efeito que cada fator tem. Existem dois tipos de efeitos: Fixo e Aleatório. O primeiro ocorre quando o pesquisador define quais serão os níveis do fator, havendo portanto, uma escolha pré-determinada. Para o efeito aleatório, é feito uma amostragem de quais serão os níveis do fator, ou seja, dentre os possíveis níveis são sorteados aqueles que farão parte do experimento.

No caso em questão, o primeiro fator P (professor) teve os níveis $i = 1, 2, 3, \dots, 10$. Os mesmos foram escolhidos aleatoriamente dentre os muitos que preenchiam os requisitos; sob este aspecto, portanto, o fator foi aleatório. O segundo fator é R (redação) e apresenta os níveis $j = 1, 2, 3, \dots, 20$; que foram amostrados do total de 30 alunos, aleatoriamente; portanto, o fator "Redação" também foi aleatório. O terceiro fator, T (Técnica), entretanto, foi fixo. Para o fator T os níveis $k = 1, 2$ foram definidos pelo pesquisador, dentre aqueles existentes. Este fato exerce influência sobre o nível de generalidade dos resultados. Se os efeitos são fixos, as conclusões ficam restritas aos tratamentos em teste, mas se os efeitos são aleatórios, as conclusões podem ser estendidas para a população de onde os tratamentos foram amostrados. Desta forma, as conclusões que aqui foram tiradas valem apenas para os professores em questão, para os alunos de Língua Inglesa Escrita III do curso de Letras da UFPR, e para as técnicas holística e analítica de correção de redações.

Além de termos testado separadamente a ação destes três fatores sobre a variável dependente, estivemos interessados nas interações entre eles, ou seja, na combinação de professor e redação (PR), professor e técnica (PT) e redação e técnica (RT), e na combinação destes três fatores simultaneamente (PRT). Para sabermos se as interações possuíam efeito fixo ou aleatório, bastou analisar os fatores. Se um dos fatores fosse aleatório, então a interação também seria de efeito aleatório; caso contrário, se todos os tratamentos que compõem a interação fossem fixos, esta também seria de efeito fixo.

Cada efeito teria um *componente de variância* (se fosse aleatório) ou um *efeito de variância* (se fosse fixo). O componente de variância é denotado pela letra que identifica o fator como índice de variância, e um efeito fixo é sempre representado pela soma de quadrados do efeito do fator dividido pelos seus graus de liberdade (ver Anexos 8, 9, 10 e 11).

Feitos todos estes cálculos, apresentamos a hipótese por eles testada:

Ho : As médias atribuídas às redações pelos professores através da técnica holística são estatisticamente iguais àquelas atribuídas pelo técnica analítica.

H1 : As médias atribuídas às redações pelos professores através da técnica holística são diferentes daquelas atribuídas pela técnica analítica.

Os resultados da análise de variância para o nosso experimento são apresentados na Tabela 10, a seguir:

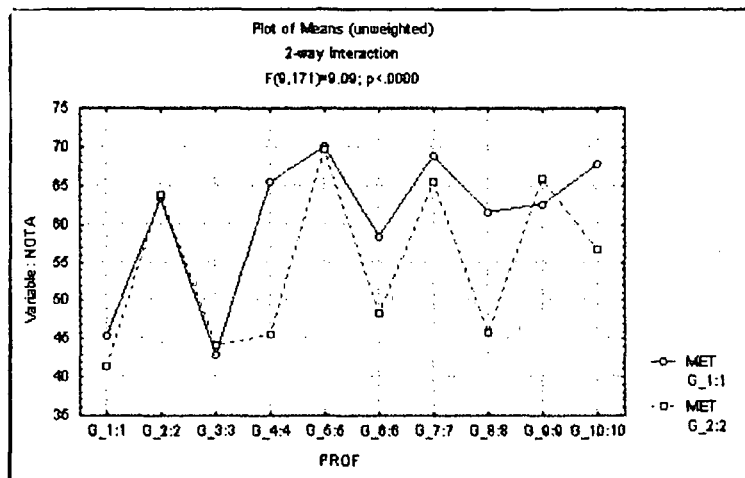
Tabela 10 : Análise de Variância para o Modelo

Fonte de Variação	SQ	g.l.	QM	Teste F	Nível de Significância
Efeitos Principais					
P: Professor	31258.587	9	3473.1763	-----	-----
R: Redação (R1)	60072.751	19	3161.7237	17.41	0.0000
T: Técnica (R2)	3552.160	1	3552.1600	4.886	0.0410
Interações					
R1= PR	31054.572	171	181.60569	-----	-----
PT (R)	5528.137	9	614.23739	9.095	0.0000
RT (R)	3426.531	19	180.34374	2.670	0.0004
R=PRT	11548.722	171	67.536389		
TOTAL (corrigido)	146441.46	399			
R2= PT+RT-PRT					

Os resultados da Tabela 10 confirmam o que já foi visto anteriormente (item 5.4), ou seja, os fatores R e T, redação e técnica respectivamente, influenciam a nota, pois o nível de significância para ambos demonstrou que a ação dos mesmos foi significativa. Contudo, o fator professor não foi passível de verificação a nível de efeito isolado sobre a nota, só aparecendo nas interações entre fatores. Quanto a estas, notamos que a um nível de significância de 5%, tanto a interação Professor/Técnica quanto Redação/Técnica foram significativas, ou seja, a técnica que atribui as maiores notas depende do professor que está corrigindo as redações e também depende da redação em

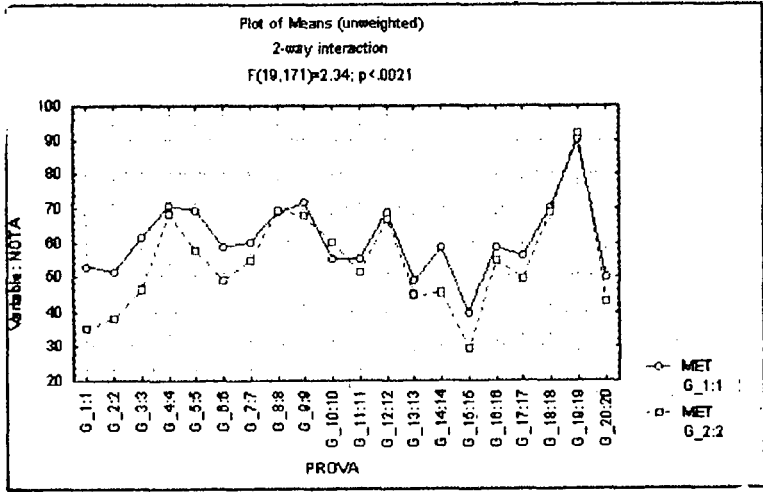
questão. Assim sendo, rejeitamos H_0 a favor da hipótese alternativa. Isto equivale a dizer que as médias obtidas pelas redações segundo a técnica analítica não são iguais, estatisticamente falando, às médias obtidas pelo técnica holística. Para saber qual dos métodos atribuiu notas maiores em termos de média, é necessário analisar professor por professor e redação por redação. Para ilustrar estas interações Professor/Técnica e Redação/Técnica às quais acabamos de nos referir, vejamos os gráficos 6 e 7.

Gráfico 6 - Interação entre Professor e Técnica



O gráfico 6 nos mostra que realmente é difícil generalizar, pois alguns professores atribuem notas mais altas quando corrigem as redações segundo a técnica holística (G1), como os Professores 1, 4, 6, 7, 8 e 10; mas outros fazem exatamente o contrário, como é o caso dos Professores 3 e 9; e há ainda casos como o dos Professores 2 e 5, que conseguem atribuir notas que convergem praticamente para a mesma média, independentemente da técnica utilizada para a correção. Isto dificulta a comparação entre as diferentes técnicas, pois nos leva a concluir muito pouco. Se o número de interações fosse menor, com certeza poderíamos tirar conclusões mais concretas.

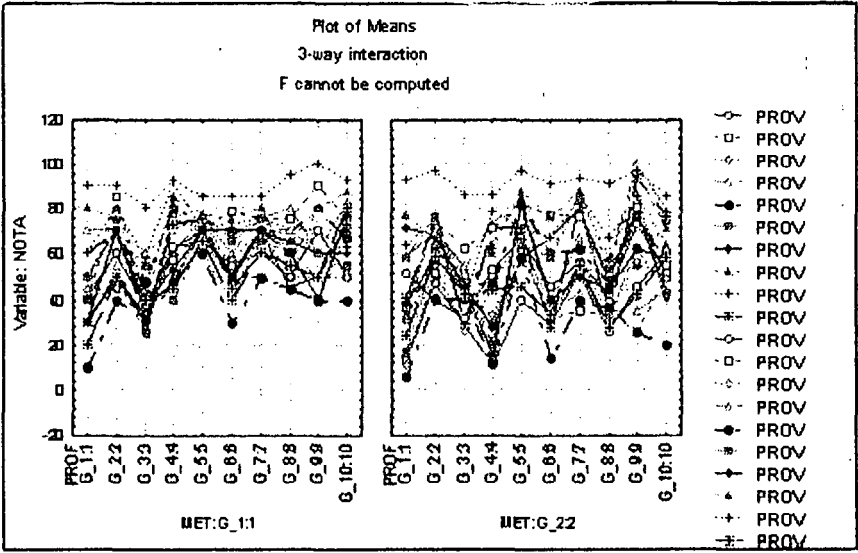
Gráfico 7 - Interação entre Redação e Técnica



No gráfico 7 podemos observar que , em geral, as médias atribuídas às redações segundo a técnica holística são maiores que àquelas atribuídas pela analítica. Este resultado nos faz inferir que quando um professor corrige uma redação seguindo critérios específicos, ele é mais rígido e encontra mais problemas a serem sanados. Talvez este fato se deva ao professor , ao avaliar um texto analiticamente, ser obrigado a observar aspectos que numa leitura global lhe passariam despercebidos por não atrapalharem o fluxo da leitura do texto.

Finalmente, vejamos no Gráfico 8 como se comportam os três fatores (redação, técnica e professor) interagindo em dois momentos diferentes. No primeiro, teremos redação, professor e técnica holística (G1) e no segundo, redação, professor e técnica analítica (G:2).

Gráfico 8 - Notas das redações (em ordem de 1 a 20), atribuídas por todos os professores



Como se pode observar, houve menor variação entre as notas atribuídas segundo a técnica holística (gráfico à esquerda) do que segundo a técnica analítica (gráfico à direita), pois, visualmente, é notório que há menos dispersão na correção segundo a técnica holística, ou seja, as linhas estão mais concentradas, formando uma faixa mais estreita que aquela da correção analítica. Esta descoberta nos surpreendeu, pois uma vez que na técnica analítica havia uma série de critérios a serem seguidos por todos os professores, esperava-se justamente que houvesse uma maior homogeneidade na sua utilização. Em outras palavras, na correção holística a nota é dada de acordo Na analítica, por outro lado, vários fatores são considerados, levando a uma maior variabilidade nas notas para os nossos professores sem treinamento. Seria interessante , em uma ocasião futura, aprofundar esta nossa pesquisa, entrevistando os professores ou aplicando um questionário com o intuito de entender melhor seu comportamento diferenciado na aplicação das duas técnicas de correção. Fica então uma sugestão para pesquisas futuras ou até mesmo para a complementação deste trabalho em uma outra oportunidade.

5.6. Considerações Finais

O cálculo dos Coeficientes de Correlação por Postos de **Spearman** para as notas atribuídas segundo as duas técnicas demonstrou resultados superiores a 0,60, ou seja, resultados que mostram uma forte relação entre as técnicas holística e analítica de correção. Além disto, os testes "t" de Student comprovaram a existência de correlação significativa entre as notas atribuídas segundo as duas diferentes técnicas. Contudo, a análise gráfica dos dados levantou a suspeita de haver interações entre as variáveis independentes, Professor, Redação e Técnica.

A partir daí, fez-se necessária a realização de uma Análise de Variância que confirmou a existência das interações Professor/Técnica e Redação/Técnica. Com isto, rejeitou-se a hipótese nula (H_0) de que as médias das notas atribuídas pelos dez professores às vinte redações pela técnica holística são iguais àquelas atribuídas pelos mesmos professores, às mesmas redações, pela técnica analítica. Praticamente este resultado quer dizer que não se pode dizer de antemão qual técnica de correção atribui maiores notas; isto dependerá do professor que está atribuindo as notas e da redação em questão. Para descobrir qual e quando uma técnica sobrepuja o outro é preciso verificar os gráficos do corpo do trabalho e dos anexos, considerando sempre cada caso em separado, nunca generalizando.

6. CONCLUSÃO

A presente pesquisa se propôs a analisar os resultados da avaliação de um conjunto de redações segundo duas técnicas distintas: a holística e a analítica. As redações foram elaboradas pelos alunos da disciplina de "Escrita III", do Curso de Graduação em Letras da UFPR, no ano letivo de 1996. Foram selecionados aleatoriamente 20 alunos para integrarem a amostra e 10 avaliadores que se enquadraram num número de requisitos previamente estabelecidos.

Os principais resultados da pesquisa foram os seguintes:

1. Há uma correlação significativa entre as técnicas de avaliação holística e analítica, ou seja, as notas atribuídas pelos 10 professores segundo cada uma das técnicas se comportam de maneira similar entre si. Mesmo assim, as médias gerais nas duas técnicas são diferentes, sendo maior na holística, o que nos permite afirmar que, segundo nossa pesquisa, os professores atribuíram notas maiores quando julgaram pela impressão global do texto.

2. A análise das médias revelou que as notas atribuídas às redações foram influenciadas pela personalidade do professor que corrigia as redações: alguns professores foram muito rigorosos na atribuição de suas notas, ficando suas médias bem abaixo das de seus colegas; outros foram muito benevolentes, mas a maioria (60%) atribuiu notas cujas médias se distribuíram de forma homogênea em torno da média geral de cada técnica de avaliação.

3. Verificou-se que, além de os fatores "Professor", "Redação" e "Técnica de Correção" terem agido isoladamente sobre a variável dependente (notas), também agiram de modo interativo. Ao observarmos a interação Professor/Técnica de Correção, constatamos que alguns professores atribuíram notas mais altas pela técnica holística, ao passo que outros atribuíram notas mais altas pela analítica e outros ainda atribuíram notas praticamente iguais na aplicação das duas técnicas de correção.

4. Constatou-se ainda uma menor dispersão das notas atribuídas segundo a técnica holística e uma maior dispersão segundo a técnica analítica.

Com base nos resultados apresentados pode-se inferir o seguinte:

•1. Embora as duas técnicas de avaliação das redações tenham levado a resultados semelhantes, a técnica holística apresenta um grau de confiabilidade maior pelo fato de revelar menor dispersão das notas em torno da média geral. Uma possível explicação para isto é que o professor, ao avaliar uma redação pela técnica holística, se deixa guiar apenas pela impressão geral, sem se ater a detalhes.

•2. A revisão bibliográfica revelou que o texto produzido constitui um todo complexo, com diferentes dimensões, tais como a sintática, a retórica e a lexical. A nota, na avaliação holística é dada pela impressão despertada no leitor em todos estes níveis. Por outro lado, na avaliação analítica, algumas destas dimensões são explicitadas, podendo ser-lhes atribuídos ainda pesos diferentes. Entretanto, mesmo a maior explicitação dos critérios de avaliação parece não conseguir, por si só, garantir uma certa homogeneidade no momento da avaliação dos textos. Como pudemos verificar através da nossa pesquisa, o mero estabelecimento prévio de critérios de avaliação de escrita não garante a ausência de dispersão dos resultados; pelo contrário, só faz acentuá-la.

•3. É preciso treinar os professores de modo intensivo e contínuo para garantir que os mesmos saibam utilizar as escalas e/ou o conjunto de critérios a serem aplicados. O treinamento dos professores garante que os critérios adotados não sejam vazios, que a dispersão das notas em torno da média seja a menor possível, evitando disparidades no julgamento de redações. É no treinamento que se estabelece quais são os "pontos zeros" e os "pontos perfeitos" (Alderson et al., 1995) de modo a assegurar que todos os professores tenham total compreensão de cada um dos critérios a serem utilizados.

•4. Além da personalidade do professor, a pesquisa realizada nos mostrou que também a técnica de correção e o próprio texto influenciam a nota. Estes são fatores que agem isoladamente ou em conjunto e podem levar a discrepâncias entre avaliadores. Certamente, o melhor modo de minimizar tais discrepâncias e atingir um maior grau de confiabilidade dos resultados da avaliação é, mais uma vez, o treinamento dos professores.

•5. A falta de treinamento não foi a única causa da instabilidade dos resultados da avaliação analítica. A falta de participação dos professores na preparação e/ou escolha da tabela de critérios também podem ter exercido uma certa influência. Embora 60% dos professores tenham conseguido uma certa homogeneidade em seu julgamento, esta proporção poderia ser maior se, após a correção holística, estes professores tivessem se reunido e organizado sua própria escala conjunta de correção.

Apresentadas as inferências resultantes da análise e interpretação dos dados, podemos afirmar que, segundo a nossa amostra, é possível confiar tanto na técnica holística quanto na técnica analítica de correção, pois ambas levaram a resultados semelhantes.

Além de nos ajudar a entender melhor o processo de avaliação de redações, este trabalho trouxe à tona algumas questões relacionadas à metodologia de avaliação da expressão escrita que podem vir a ser objeto de pesquisas futuras. Por exemplo, seria possível refazer a pesquisa utilizando as mesmas redações e os mesmos professores, só que agora treinando estes professores de modo a averiguar se há realmente a minimização das disparidades por nós detectadas. Esta nova pesquisa, com certeza, levaria a um resultado mais incisivo a respeito da interação Professor/ Técnica. Uma outra sugestão de cunho metodológico, seria entrevistar os professores imediatamente após a correção holística e tentar descobrir porquê eles atribuíram aquelas notas; ou seja, tentar definir que critérios internos cada professor utilizou para dar notas segundo a impressão global do texto.

Tais sugestões enriqueceriam muito o nosso trabalho com relação à prática de ensino e nos ajudariam a entender melhor os caminhos do processo de avaliação da escrita. Além destas sugestões de extensão de nossa pesquisa, há algumas outras de ordem prática que gostaríamos de ressaltar. Em primeiro lugar, é importante que se compare "coisas que sejam comparáveis" entre si, desde o princípio, para evitar que os dados tenham que ser transformados (ver Capítulo 5). Isto poupará tempo ao pesquisador que poderá se dedicar com maior afinho à análise e interpretação dos dados e não à sua manipulação.

A segunda sugestão seria delimitar o campo da pesquisa de modo a abranger disciplinas relacionadas ao mesmo e que sejam de algum domínio do pesquisador; caso contrário, o pesquisador acabará tendo muito trabalho extra e correndo o risco de se perder no caminho.

A terceira sugestão de ordem prática diz respeito à coleta de dados. Num trabalho como o nosso em que foram utilizados o mesmo conjunto de redações e o mesmo grupo de professores em dois momentos diferentes, seria interessante que o intervalo de tempo entre as avaliações fosse maior. Um intervalo de no mínimo 24 horas e no máximo 10 dias entre a avaliação holística e a analítica fez com que esta variável ficasse fora do nosso controle total, o que obscurece um pouco os resultados. O ideal teria sido a reunião de todos os professores fazendo as avaliações com o tempo controlado para as duas técnicas, de forma que a variável tempo estivesse totalmente sob controle. Além disso, o intervalo ideal entre uma avaliação e outra deveria ser de, no mínimo, 30 dias para assegurar que os professores não se lembrassem de fato das notas atribuídas na primeira fase. Não pudemos trabalhar o fator tempo desta maneira por restrição do nosso cronograma e devido à dificuldade de reunir os 10 professores. Além disso, cometemos o erro de informar os professores, na segunda fase, que eles estavam corrigindo redações que já haviam corrigido anteriormente (ver Anexo 2). Segundo Alderson et al. (1995) deve-se sonhar ao avaliador este tipo de informação, para evitar que o mesmo se deixe influenciar pelo seu próprio julgamento prévio e para podermos verificar se os avaliadores apresentam consistência em suas avaliações.

Como nossa pesquisa foi centrada na avaliação da expressão escrita, não poderíamos deixar de fazer também algumas sugestões para a prática do ensino. Entretanto, as mesmas são relacionadas à área de avaliação de escrita e não ao ensino propriamente dito.

A primeira sugestão é a de que a elaboração da escala, holística ou analítica, a ser utilizada deve ir de encontro aos objetivos da disciplina e ser elaborada de modo a atender tais objetivos. Uma vez escolhida a forma de avaliação e elaborada a escala de avaliação a ser adotada, os professores devem ser treinados para a sua utilização, visto que constatamos que de nada adianta uma escala de critérios sem o treinamento prévio dos avaliadores que dela farão uso. É importante ressaltar que todos os professores que ensinam ou possam vir a ensinar num curso que se utiliza de uma escala

elaborada ou adotada precisam ser treinados de modo a garantir a consistência de sua avaliação e a confiabilidade do sistema de avaliação da disciplina como um todo. Professores despreparados ou mal preparados mudarão seus padrões de avaliação durante o processo, usarão a escala de critérios de forma inconsistente e/ou inadequada, e terão dificuldades para adequarem sua avaliação à de outros professores. Segundo Alderson et al. (1995) um programa efetivo de monitoramento atenua todos estes agravantes e ajuda a desenvolver o senso de trabalho em equipe com a presença de um líder, que no caso poderia ser o coordenador de área.

Seja a escala holística ou analítica, é imprescindível que a avaliação escrita seja feita por dois professores, no mínimo. Cada um fará a sua avaliação numa folha em separado do texto em julgamento para evitar, assim, que um professor influencie ou se deixe influenciar pela avaliação do colega. A nota final do texto será a média aritmética das notas atribuídas pelos professores. Alderson et al. (1995) afirmam que, em termos práticos, dois professores são suficientes, e no caso de total desacordo entre eles entra em cena a figura do monitor, coordenador ou líder para resolver o impasse. A solução mais comum, entretanto, seria a de chamar um terceiro professor para avaliar o texto em questão e fazer a média entre as duas notas mais próximas, desprezando-se assim aquela distoante. Hughes (1994) por sua vez, frisa que com um número superior a quatro professores avaliando o mesmo texto, o processo avaliativo ganha muito pouco em termos de confiabilidade; ou seja, não apresentará resultados significativamente diferentes dos apresentados por quatro avaliadores. Todavia, se por um lado o processo avaliativo ganha em confiabilidade ao termos quatro professores corrigindo a mesma redação, por outro lado ele se torna pouco prático, pois é difícil envolver quatro professores de uma mesma instituição na avaliação de testes corriqueiros de escrita.

Outra sugestão prática diz respeito à avaliação analítica. Segundo autores como Weir (1993) e Hamp-Lyons (1991) este tipo de avaliação tem uma função diagnóstica, que facilita o preparo de aulas e o acompanhamento do progresso dos alunos, uma vez que ela capta as fraquezas e os pontos fortes dos alunos. Desta forma, ela pode ser usada no começo de um curso como o de Língua Escrita III para auxiliar os professores no desenvolvimento de um programa específico para atender às necessidades de seus alunos e, no decorrer do curso, para verificar os progressos do grupo.

Este tipo de avaliação dá ao professor uma medida exata das dificuldades que seus alunos têm, e ajuda a organizar os meios para sanar tais problemas, individualmente ou em grupo.

Finalmente, nossa última sugestão seria a de informar os alunos sobre os critérios de avaliação antes dos mesmos serem aplicados. É importante que o aluno saiba exatamente como será avaliado e tenha exemplos de textos considerados bons e ruins pelo professor e pela respectiva instituição de ensino. Desta maneira, o aluno estará seguro de como deve proceder ao produzir um texto para avaliação, e não se sentirá desvalorizado ou inferiorizado pela mesma. Este pequeno detalhe faz uma grande diferença, pois evita aborrecimentos futuros para o aluno e para o professor, uma vez que cada um sabe qual é o seu papel e o que deles se espera, não havendo surpresas desagradáveis. Por exemplo, o aluno não produzirá um texto rico em estruturas complexas e criativas se souber que será penalizado por não tê-las organizado de forma perfeita. Pelo contrário, se ele sabe que pode ser penalizado por isto, procurará transmitir sua mensagem através de frases curtas e de estruturas simples que lhes sejam familiares, diminuindo assim as chances de ser penalizado por se arriscar demais. O professor, por outro lado, não terá dificuldades para se explicar ao lhe ser solicitada uma revisão de provas ou até mesmo uma justificativa de reprovação, em caso extremo, pois estará amparado por um conjunto de critérios claros e do conhecimento de todos.

Com estas sugestões de cunho prático-metodológico do exercício da docência, encerramos o nosso trabalho com a certeza de que o mesmo trouxe à luz algumas questões relevantes para a metodologia da avaliação da escrita e contribuiu para uma reflexão crítica da mesma.

ANEXO 1 - AS REDAÇÕES

Redação 1

The "Vestibular"

I agree with the form of admission into a Brazilian university called "vestibular" but not totally agree because there are many things that I think wrong. The idea about the test is good due to the evaluation that it does to choose only students who are able to.

The entrance examination is an interesting form to guide the students to study a lot of previous subjects and to get necessary knowledge to keep on your course at the university. However, one of the things that is not correct in my opinion are the matters required in areas completely different. For example, when it is required physics for the Letter course. For me this is not correct. I think that the matters must be according to the courses. Another thing among the most important would be not to make the examination after the high school but it could be made at the last year of the high school. Then the high school should prepare better the students so that they were able to make the test in that time.

Redação 2

The "vestibular"

For Brazilian students it is the only test that gives them chance to enter Brazilian Universities. The entrance examination it is not easy. The "vestibulandos" have to study hard several different subjects such as Chemistry, History, Portuguese, Mathematics, Physics, and others in order to pass in the tests. For most students all these subjects are not fair. They think, for example, if the "vestibulando" intends to take an English course he or she doesn't have to know about Mathematics or Physics. Probably they won't use its during the course. On the other hand, teachers and others people who work during the "vestibular" think it is the only way for the "vestibulandos" enter University well prepared.

Redação 3

The Vestibular

It's particularly difficult for me to have a finished opinion about the "Vestibular" because I have to consider lots of things; for example, there are lots of students who want to study in an University (specially Federal University); there are not so much place as it is necessary; the students are not well-prepared, in other words, they don't have sufficient knowledge. On the other hand, the students who can pay a course which teaches how to give correct answers have more opportunities than the others who can not pay it. This is a simplified way to consider the problems of the "Vestibular".

When I look all these I think that the "Vestibular" is not a good thing but I can't see another way to test the students on different subjects. It should be necessary to change the Educational Brazilian System to eliminate the "Vestibular".

Redação 4

The Vestibular

Entrance exams is the only form of admission into a Brazilian university. They adopt this method to select the students because there are many people applying for few places in the universities, mainly the public ones. It means that in Brazil college can be free of charges.

However, to get into a public university they must study very hard. They consider vestibular as a "real nightmare" because they are supposed to know all contents that they learnt in High School.

Most between them, before getting into the college, attend a preparatory course, which gives them conditions to revise many contents that they haven't learnt or forgot. These preparatory courses are really very expensive, most times as expensive as a course in a private university. So they rather pay these courses for a year than five years in a private college.

There is an enormous polemic around vestibular, many people agree that it should be abolished and the students would be evaluated during their school life but probably there wouldn't be places for everybody. Then, while they don't find another effective method to select the candidates, Brazilian students will face "vestibular" every year.

Redação 5

"The Vestibular"

The "Vestibular" is something like a ghost. It tortures the students for one year or more during their hard preparation for entrance examination. The preparation to the exam consists of all subjects learned during secondary and on the end of the year when all students are exhausted and stressed, the "vestibular" starts and remains by a period of four endless days. There is a great competition and only well prepared students will have chance of success.

A new kind of evaluation should be adopted in order to accept students on Universities, for example, only one test could be done each year by government and based on their grades students would be accepted on this or that university. Another possibility would be a specific test linked to the area that students have chosen, i.e. if a student wants to study medicine only a biology test and plus language should be required. Why a candidate for a law course must to know about chemistry or mathematics?

Redação 6

The "Vestibular"

Vestibular is a word that makes some people specially that ones who already finished their secondary school, feel panic, depression, and other feelings no-one can imagine. This word means the promise of a good life by studying.

Although all this illusion of glory, many people will never get it because in our country this exam is the only way to get into a university. But it is really difficult to imagine how a person can show everything he/she knows in only one test. Public universities are suppose to attend people that can not pay for a private one, but it is not the reality.

Many people that frequent public universities are these ones who had money to pay for a good preparatory school. In order to change this picture , this exam would be changed for other way of evaluation. The government could determine a comission to evaluate people since the beggining, in the first step in a school and give conditions os studies as good classes, teachers for every class they built. Maybe when this dream comes true our country will grow to a Big Nation where there is no powerty, neither illiterate people.

Redação 7

The vestibular

In my opinion, vestibular has a good and negative side. The good one is that vestibular is a way to select students which are prepared to start a University course. The students are evaluate across their knowledge about a determined subject and for the time being, this is the only form of Brazilian University.

On the other hand, this entrance examination has some negative consequences on students feelings. They have a big emotional tear before and during the tests. There are many cases of students that although well prepared for the tests, become very nervous and as a consequence they obtain bad results.

Besides some students can't do a prepared course for vestibular after their high school, because nowadays this kind of course are so expensive. However, I can't see another form to entrance to the university. Maybe one day, Vestibular could be abolished and high schools can develop more qualiffied course for their students in order to the students go directly to the university after finished high school. But now, Vestibular is inevitable.

Redação 8

The "vestibular"

I think that the "Vestibular" is not a good form to evaluate students for three reasons. First of all, here in Brazil we have many differences between public and private schools in terms of instruction. For example, private schools give more material conditions such as good rooms and laboratories than public schools. Furthermore, in private schools the teachers teach all the subjects which are demanded in "Vestibular" while in public schools it is not possible because we usually have a strike a year. The second reason is that rich people have more chances in this exam than poor people because they not only study in private schools but can also pay the "cursinhos". The third reason is that in this exam there are many subjects in a single question which can make the students to be nervous.

In summary I think that for these reasons the "Vestibular" is not a good form to evaluate students knowledge and we should create another kind of entrance examination,

Redação 9

The "vestibular"

The entrance examination in Brazil, which is called "vestibular", is often too demanding and unfair. First, there are too many students trying to get into the universities but there aren't enough universities. Besides, most of the time just private schools offer the students a satisfactory basis to take the "vestibular" tests and succeed. However, good private schools are extremely expensive. Second, even though someone has studied in a good private school he/she will have to obtain a good grade in all disciplines even the ones that are useless for his/her area of interest, to enter.

It must be explained that these problems occur mainly when getting in public universities. So, there are chances on passing on the private universities tests but then, the problem is that many people can't afford it and they are forced to give up. It can be concluded that taking a complete University course in Brazil is just for a small number of privileged people.

Redação 10

Vestibular

Vestibular is the most important test for all teenagers between 17 and 20 years old who intend to study in a university. It is the only form of admission into a Brazilian university. Because of it, many people believe that vestibular isn't the appropriate form to select students. They believe that there are another ways to do this selection, such as, the grade that students took in high school.

However, I disagree with it. I believe that vestibular is, nowadays, the best way to select students for universities. It is impossible a student who do just a "supletivo" come into a university at the same way that a student who had to study hardly in a select school.

In conclusion, vestibular must be the only way of admission into a Brazilian university, because it is the most appropriate form to select students.

Redação 11

(O aluno não colocou título)

It is not fair to every student how they should enter the university in Brazil, since the only way is by an entrance examination. The principal reason of this injustice is money. If the person belongs to an ordinary family, probably he will study in a public school. Everybody knows the level of knowledge the student gets. Because of low salaries teachers, sometimes are not motivated to teach. Moreover, there are schools without a good infrastructure. Many times people do not have where to sit down. Also we can find students that have to work to support themselves, and the school, what is the most important is left behind.

These things do not happen in private schools. Every student has a better system of studying. The contents are stronger and better taught. After that, when near the examination this lucky student pays a course that gives all support necessary to get into the university. So, how can a poor person pass all the tests? Studying to death? What are the chances of success ? Very few ones!

Redação 12

Vestibular

One of terrible situation that students need to face when they graduate in high school and intend to be admitted in University is to pass in a examination called vestibular.

The exam called vestibular has created a lot of controversy among people. Some experts claim that students have to do special courses before vestibular in order to improve their knowledge but in my own opinion this kind of test does not evaluate the real knowledge of students. The right way to accept students in the university would be through high school examinations. The student would be evaluated by all grades he/she had in his/her course. The major problem that people feel is that during the exams, they shows clearly symptoms of stress, lack of memory and concentration and others psychological disturbances creating an hazardous situation for them.

But unfortunately Brazil is a poor country and it is necessary to have this kind of exam because universities offer few opprotunities to students and they must be selected.

Redação 13

"Vestibular"

Brazilian students have to study very hard in order to pass in "Vestibular", to get into the university.

There are lots of "cursinhos" specially for this case, that force students to choose one of them in order to give the knowledge they need to pass. But, if the schools were better, students wouldn't need study at "cursinhos".

Although this kind of entrance examination, the system doesn't have worked well and has occurred some injustices with many students that fail.

Therefore, the system should change. There are many ways to examine students to the University. One of them could be an evaluation in the grades since the first school of each student, and then the student could choose a course according to the grades.

Redação 14

"Vestibular"

Vestibular has become a real "gold eggs chicken" in Brazil. Today, so important as to study during all your schoolar life in a good school is to be in a good "cursinho" at the end of high school. There is no more books, students learn (?) with "apostilas" and spend all high school time memorizing formulas and summaries. Perhaps the huge number of zeros in the last vestibular could be a proof of what is going on. Shall we blame against the "cursinhos", whose owners are collecting the golden eggs? I don't think so.

Of course that "cursinhos" are not the guilties. They just offer what "vetibular" wants from pupils: "X" in the right answer. The matter is that, in a destroyed educational system as Brazilian one, nothing else seems to be available. The most part of students don't know even write a simple text. Asking for wrote answers would empty the University, what would be certainly worse. This situation comes out as result of cultural- economic situation: no good schools, nor public or private, that work under an universal need for technical specialization, in despite of a humanitarian knowledge.

Redação 15

"Vestibular"

In my opinion, Vestibular isn't a practical way to enter in an university, only a low percent can pay a previous course. Some schools ignore the fact, others not. The annual curriculum is imposed by the principal and Vestibular rules are forgotten. Then, students have gone out to the school in order to find a specific course where they will get all the "tricks" to pass in a Vestibular. But, these courses are so expensive that students chances decrease. If you wish study in an university, try to get good grades before thinking enter in an university. Later on, public or private schools won't prepare students to Vestibular. Schools will prepare students to obtain professional experience because Brazil needs technical knowledge.

Redação 16

"Vestibular"

Vestibular it's a word which sounds a terrible thing for all the students who want to enter in the University. It looks like a monster that you need to fight against. And why?

Firstly, because students usually make tthe exam under pressure because they think that they have obligation to pass.

Secondly, because they do the exam stressed because they send their free time , such as weekends and even nights, studying a lot.

Finally, because sometimes the result of the exam is unfair because when students failed they feel themselves like a loser.

In my opinion, the admission into a university would be automaticaly and the student would choose what the university he wants. But, on the other hand, the University can say "yes"or "no" to this student based on student's grades.

The entrance examination

The entrance examination has been adopted by several countries, even by for those classified as belonging to developed nations, such as France, Swiss and some other european institutions considered as well.

In Brazil, the educational system in force is so old as the republic itself. Despite the fact the number of vacancies available afforded in college institutions remains still at power levels, applicants numbers raise at each year. But, they cannot step in at once, what lead many to stay apart from oportunities. Between the main reasons are: non affording conditions, incompatible schedules.

The entrance exam is far away from being the utmost solution for the problem, but beside this fact, other options such as the American model can coexist at the same time between us and good results can be obtained by creating maybe a third system that could rally the advantages of both.

Redação 18

"The Vestibular"

For many students, the hardest part of their student's life is when they have to deal with "the vestibular". If they want to hold on to their studies, they have to do these tests, otherwise they won't be able to study in a Brazilian university.

Many people consider it unreasonable. You can be a great student, study a lot, but if at the exactly moment of the test you get nervous, you'll probably fail, because you couldn't concentrate yourself. Then you may feel the worst person in the world.

On the other hand, "the vestibular" is the only way the universities have to control the number of students that will be admitted on it, because there are not enough vacancies for every student.

"The Vestibular" is the only way, nowadays, universities have to pick up the students. When the government decides to change the way it is done, probably the good students who couldn't be admitted on the universities will have a chance.

The Vestibular

It is usually in January when most of the former High School students in Brazil have to face one of the hardest tests of their lives: the so called Vestibular.

The Vestibular is normally a four day test where the student is supposed to show how much he knows on almost all the subjects he has studied so far. This definition of the Vestibular may sound odd to someone who comes from a country where the entire school life of the student is taken into consideration before the Institution accepts or refuses him as an University student. But, anyway, this is the way things work in Brazil. If you have been an excellent student for your whole life, simply adore Medicine but unfortunately did not do good enough on the Physics test, you will have to wait one more year (or maybe more!) to get into the Medical University.

Because of this test being so unfair, many people think it should be no longer used. But the point is, if the Vestibular becomes something from the past, how else will the selection be done?

It is an Utopia to think of using the same system as developed countries. They have excellent public schools where their students finish high school as well prepared as any private school student. Comparing grades of a student of a Public school in Brazil to the ones of a private school and by then choosing one to enter an University would be as unfair as any stressing Vestibular.

In this way, Improving the school system in Brazil would be the best solution to finish with the Vestibular and make the entrance to an University much more fair and pleasant to the students.

Redação 20

Vestibular

The entrance examination to Brazilian University isn't fair sometimes. There are many problems that damage the opportunity to students make a university course.

Rich students generally study in private schools , they have a better education from the others who study in public schools. Furthermore, some of the poor students must work too, and they don't have time to be dedicated students. The method used at examen don't exclude the possibility to be aprooved by luck though, it is more difficult. On the other hand, if there aren;t an entrance examen how to choose the good students? To choose by their curriculum isn't fair too, all people have chance to improove themselves sometime in life. Furthermore could be easy to rich people pay for good grades. So entrance examination can't be fair sometimes, but it is necessary.

ANEXO 2- INSTRUÇÕES PARA OS PROFESSORES

Procedures for evaluation

FIRST PART

HOLISTIC OR GENERAL IMPRESSION MARKING

You are a teacher scoring class compositions on your own. Read each composition quickly for an overall first impression, paying particular attention to the *message* which the writer is trying to get across. Once you have finished reading the compositions and have a standard already established in your mind, reread each composition and give it a grade ranging from 0 to 100 bearing in mind that 70 would be the pass mark.

As soon as you have graded your compositions, do not read them again or change the grades, just put them back in the envelope.

Wait at least 24 hours before you go through the second part of this project.

Information you need in order to grade the compositions holistically:

1. Subjects: undergraduate students of the Writing III course in their first week of academic activities
2. Level of students: in order to pass Writing III a student must be at First Certificate level
3. Kind of writing: academic
4. Function: argumentative
5. Objective: to express personal opinion
6. Audience: the writing teacher
7. Task: give your opinion about the entrance examination which is the only form of admission into a Brazilian university.
8. Time limit: 40 minutes
9. Length limit: 20 lines
10. Purpose of the task: help organize the learning/teaching material, help diagnose student's strengths and weaknesses in the language in order to follow their progress throughout the course.

So far, thank you so much for your help.

Procedures for Evaluation

SECOND PART

ANALYTICAL MARKING

Here you have the same set of compositions you already graded a while ago. Now you are going to use a profile with specifically established criteria to give them grades.

Steps to follow:

1. Read the profile in order to understand what you are about to do. Notice that it contains 5 component scales, each focusing on an important aspect of composition.
2. Read the composition quickly for an overall first impression, paying particular attention to the *message* which the writer is trying to get across. *Trust your judgment.*
3. In the *Content* and *Organization* components of the Profile, find the descriptors that best describe the writer's success in delivering a message. Determine a score for each component to reflect these descriptors and record the scores in the spaces provided (not that the descriptors for a level have to apply in order to score a composition within a level. Sometimes you may find that none of the descriptors is exactly appropriate; in such cases, choose one of the more general value terms - GOOD TO AVERAGE - and then determine the number best associated with that quality).
4. Quickly reread the composition and record scores in the remaining three components after identifying the appropriate criteria descriptors.
5. Put the compositions back in the envelope.

Once again, thank you so much for your help, I really appreciate your dedication and consideration.

ESL COMPOSITION PROFILE

CONTENT	Level	Criteria	Score
	30-27	EXCELLENT TO VERY GOOD: knowledgeable, substantive, thorough development of thesis, relevant to assigned topic	
	26-22	GOOD TO AVERAGE: some knowledge of subject, adequate range, limited development of thesis, mostly relevant to topic, but lacks detail	
	21-17	FAIR TO POOR: limited knowledge of subject, little substance, inadequate development of topic	
	16-13	VERY POOR: does not show knowledge of subject, non substantive, not pertinent or not enough to evaluate	
ORGANIZATION			
	20-18	EXCELLENT TO VERY GOOD: fluent expression, ideas clearly stated/supported, succinct, well-organized, logical sequencing, cohesive	
	17-14	GOOD TO AVERAGE: somewhat choppy, loosely organized but main ideas stand out, limited support, logical but incomplete sequencing	
	13-10	FAIR TO POOR: non-fluent, ideas confused or disconnected, lacks logical sequencing and development	
	9-7	VERY POOR: does not communicate, no organization or not enough to evaluate	
VOCABULARY			
	20-18	EXCELLENT TO VERY GOOD: sophisticated range, effective word/idiom choice and usage, word form mastery, appropriate register	
	17-14	GOOD TO AVERAGE: adequate range, occasional errors of word/idiom form, choice, usage but meaning not obscured	
	13-10	FAIR TO POOR: limited range, frequent errors of word/idiom form, choice, usage, meaning confused or obscured	
	9-7	VERY POOR: essentially translation, little knowledge of English vocabulary, idioms, word form or not enough to evaluate	
LANGUAGE USE			
	25-22	EXCELLENT TO VERY GOOD: effective complex constructions, few errors of agreement, tense, number, word order/function, articles, pronouns, prepositions	

	21-18	GOOD TO AVERAGE: effective but simple constructions, minor problems in complex constructions, several errors of agreement, tense, number, word order/function, articles, pronouns, prepositions, but meaning seldom obscure	
	17-11	FAIR TO POOR: major problems in simple/complex constructions, frequent errors of negation, agreement, tense, number, word order/function, articles, pronouns, prepositions and/or fragments, run-ons, deletions, meaning confused and obscured	
	10-5	VERY POOR: virtually no mastery of sentence construction rules, dominated by errors, does not communicate or not enough to evaluate	
MECHANICS			
	5	EXCELLENT TO VERY GOOD: demonstrates mastery of conventions, few errors of spelling, punctuation, capitalization, paragraphing	
	4	GOOD TO AVERAGE: occasional errors of spelling, punctuation, capitalization, paragraphing but meaning not obscured	
	3	FAIR TO POOR: frequent errors of spelling, punctuation, capitalization, paragraphing, meaning confused or obscured	
	2	VERY POOR: no mastery of conventions, dominated by errors of spelling, punctuation, capitalization, paragraphing or not enough to evaluate	
			Total Score:

Jacobs, et al., 1983.

ANEXO 3 - ESCALA DE CRITÉRIOS DENOMINADA *PMP2*

"New Profile Scale and Profile Method 2 (PMP2)"

Escore	Qualidade Comunicativa	Organização	Argumentação	Acuidade Lingüística	Adequação Lingüística
9	O texto demonstra uma habilidade comunicativa que dá ao leitor satisfação total.	O texto demonstra uma estrutura completamente organizada e lógica, que permite que a mensagem seja transmitida sem dificuldades.	Argumentos relevantes são apresentados de um modo interessante, com idéias principais clara e destacadamente colocadas, com material de apoio pertinente; os argumentos são efetivamente relacionados às experiências e pontos de vista do autor.	O leitor não detecta erros de vocabulário, ortografia, pontuação ou gramática.	Há habilidade total na manipulação adequada do sistema linguístico.
8	O texto demonstra uma habilidade comunicativa que não traz dificuldades ao leitor.	O texto demonstra uma estrutura organizada e lógica que permite que a mensagem seja facilmente transmitida.	Argumentos relevantes são apresentados de um modo interessante, com idéias principais destacadas e com material de apoio pertinente; os argumentos são bem relacionados às experiências e pontos de vista do autor.	O leitor não detecta erros significantes de vocabulário, ortografia, pontuação ou gramática.	Habilidade na manipulação adequada do sistema linguístico.
7	O texto demonstra uma habilidade comunicativa que traz poucas dificuldades ao leitor.	O texto demonstra uma boa estrutura organizacional que permite que a mensagem seja transmitida sem muito esforço.	Argumentos são bem apresentados, com material de apoio pertinente e uma tentativa de relacioná-los às experiências e pontos de vista do autor.	O leitor está ciente, mas não incomodado com os eventuais erros mínimos de vocabulário, ortografia, pontuação ou gramática.	Há limitações menores na habilidade de manipular adequadamente o sistema linguístico, mas que não atrapalham o leitor.

6	O texto demonstra habilidade comunicativa, embora apresente uma certa tensão para o leitor.	O texto demonstra uma estrutura suficientemente organizada para permitir a transmissão da mensagem.	Argumentos são apresentados, embora possa ser difícil para o leitor distinguir as idéias principais das de apoio, sendo que as idéias principais podem ter ficado sem suporte; sua importância pode ser duvidosa e os argumentos não parecem poder ser relacionados às experiências ou pontos de vista do autor.	O leitor está ciente dos eventuais erros de vocabulário, ortografia, pontuação ou gramática.	Há limitações na habilidade de manipular adequadamente o sistema lingüístico, embora isto apenas, eventualmente, constitua um empecilho.
5	O texto demonstra habilidade comunicativa, embora apresente uma frequente tensão para o leitor.	O texto demonstra uma estrutura organizada o suficiente para permitir que a mensagem seja transmitida no decorrer da maior parte do texto.	Argumentos são apresentados, embora não sejam relevantes, claros, consistentes ou desenvolvidos; os mesmos podem ser relacionados às experiências ou pontos de vista do autor.	O leitor está ciente dos erros de vocabulário, ortografia, pontuação ou gramática que frequentemente constituem um empecilho para o leitor.	Há limitações na habilidade de manipular adequadamente o sistema lingüístico, que frequentemente constituem um empecilho.
4	O texto demonstra habilidade comunicativa limitada, que frequentemente propicia tensão ao leitor.	O texto não demonstra uma estrutura organizada e a mensagem é difícil de ser seguida.	Os argumentos são inadequadamente apresentados e desenvolvidos. As experiências ou pontos de vista do autor são difíceis de serem detectados.	O leitor considera o domínio de vocabulário, ortografia, pontuação e gramática inadequado.	Inabilidade ao manipular adequadamente o sistema lingüístico, que causa uma grande tensão no leitor.
3	O texto demonstra habilidade comunicativa limitada, embora o significado possa eventualmente vir à tona.	O texto não revela uma estrutura organizada que possa ser detectada e a mensagem não é passível de acompanhamento.	Algumas informações são apresentadas, mas o leitor não é provido de nenhum argumento ou o argumento é basicamente irrelevante.	O leitor tem consciência dos erros grosseiros de vocabulário, ortografia, pontuação e gramática.	Há um mínimo ou nenhum senso de adequação lingüística, embora haja evidências de estruturas frasais.
2	O texto não apresenta habilidade comunicativa.	Não há estrutura organizacional ou mensagem que possam ser detectadas.	Há evidências ocasionais de algum conteúdo, mas que não é relevante.	O leitor não percebe evidências de domínio de vocabulário, ortografia, pontuação ou gramática.	Não há sinal de adequação lingüística.

ANEXO 4 - ESCALA DE CRITÉRIOS PARA O *IELTS*

PROFILE AND DESCRIPTORS - ACADEMIC AND GENERAL TRAINING WRITING MODULES TASK 2

Communicative Quality	Arguments, Ideas & Evidence	Vocabulary & Sentence Structure
9. The reader finds the writing completely satisfactory.	A clear point of view is presented and developed. The argument proceeds logically through the text, with a very clear progression of ideas. There is plentiful material.	A wide range of vocabulary and sentence structures is used accurately and appropriately.
8. The reader finds the writing communicates fluently.	A clear point of view is presented and developed. The argument proceeds logically through the text, with a very clear progression of ideas.	The range of vocabulary and sentence structures used is good, and well controlled for accuracy and appropriacy. There are no significant errors in word formation or spelling.
7. The reader finds the writing satisfactory and that it generally communicates fluently with only occasional lapses.	A generally clear point of view is presented. The argument has a clear progression overall, and ideas and evidence are relevant and sufficient, although there may be minor isolated problems in these areas.	A satisfactory vocabulary and sentence structure occurs, usually used appropriately. There are only occasional minor flaws in word formation and in control of sentence structure. Spelling errors may occur, but they are not intrusive.
6. The reader finds the writing mainly satisfactory and that it communicates with some degree of fluency. Although there is occasional strain for the reader, control of organisational patterns and devices is evident.	A point of view is presented, although it may become unclear in places. The progression of the argument is generally clear. The relevance of some ideas or evidence may be dubious and more specific support may seem desirable.	Vocabulary and sentence structures are generally adequate and appropriate, but the reader may feel that control is achieved through the use of a restricted range. In contrast, examples of the use of a wider range of sentence structures are not marked by the same level of accuracy. Some errors in word choice, word formation and spelling may occur, but they are only slightly intrusive.
5. The writing sometimes causes strain for the reader. While the reader is aware of an overall lack of fluency, there is a sense of an answer which has an underlying coherence.	The writing introduces ideas, although they may be limited in number or insufficiently developed. A point of view may be evident, but arguments may lack clarity, relevance, consistency or support.	The range of vocabulary and the appropriacy of its use are limited. There is a limited range of sentence structures and the greatest accuracy is achieved on short, simple sentences. Inappropriate choice of words and errors in areas such as agreement of tenses or subject/verb agreement are noticeable. Word formation and spelling errors may be quite intrusive.
4. The writing attempts communication but the meaning may come through only after considerable effort by the reader.	There are signs of a point of view, but main ideas are difficult to distinguish from supporting material and the amount of support is inadequate. Such evidence and ideas as are presented may not be relevant. There is no clear progression to the argument.	The use of vocabulary is often inadequate and/or inappropriate and limited control of sentence structures, even short, simple ones, is evident. Choice of words can cause significant problems for the reader. Errors in such areas as agreement of tenses or subject/verb agreement, word formation and spelling can cause severe strain for the reader.

3. The seriousness of the problems in the writing prevents meaning from coming through more than spasmodically.	The writing has few ideas and no apparent development. Such evidence ideas as are presented are largely irrelevant. There is little comprehensible point of view.	Control of vocabulary and sentence structures is evident only occasionally and errors predominate.
2. The writing displays almost no ability to communicate.	There may be glimpses of one or two ideas without development.	There is little or no evidence of control of sentence structure, vocabulary, word form or spelling.
1. The writing appears to be by a virtual non-writer, containing no assessable strings of English writing. Answers of less than two lines are automatically scored as Band 1		
0. Should only be used where a candidate did not attend or attempt the questio in any way.		

UNIVERSITY of CAMBRIDGE LOCAL EXAMINATION SYNDICATE, 1995: 101.

**ANEXO 5 - GRÁFICOS DAS NOTAS ATRIBUÍDAS PELA TÉCNICA HOLÍSTICA E
ANALÍTICA PARA OS PROFESSORES 2, 3, 4, 5, 6, 7, 8, 9 E 10**

Gráfico 1 - Notas pela Técnica Holística e analítica para o Professor 2

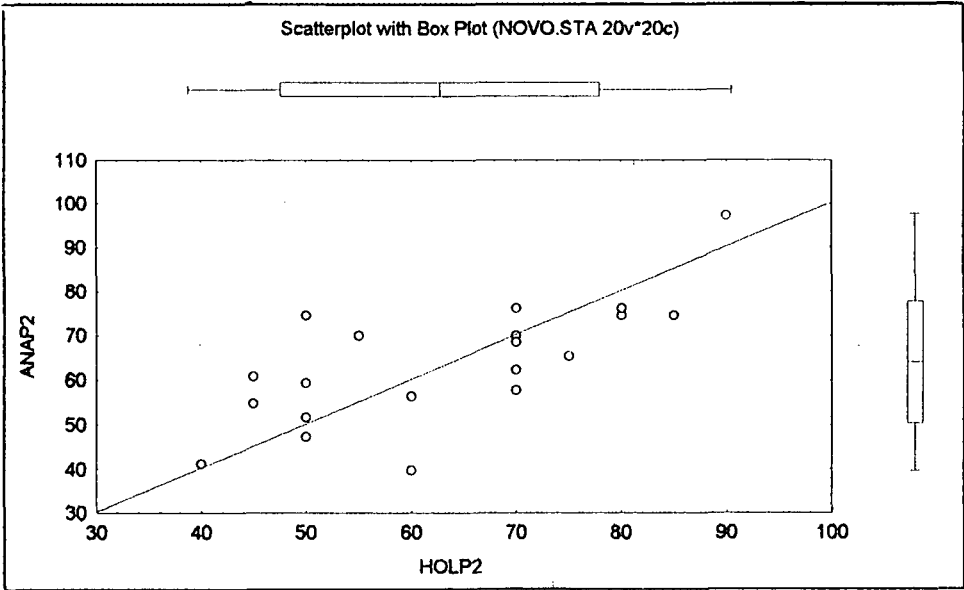


Gráfico 2 - Notas pela Técnica Holística e analítica para o Professor 3

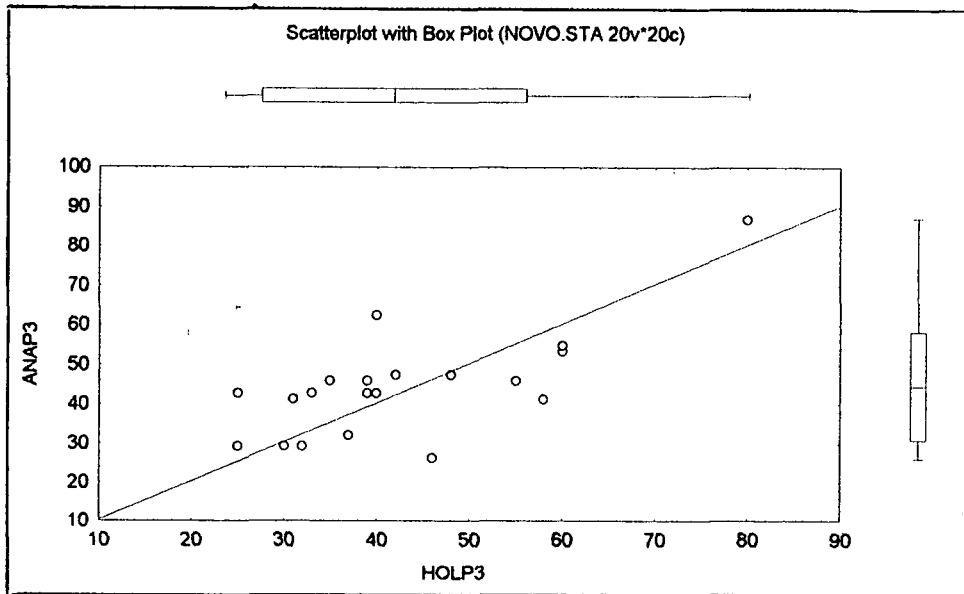


Gráfico 3 - Notas pela Técnica Holística e analítica para o Professor 4

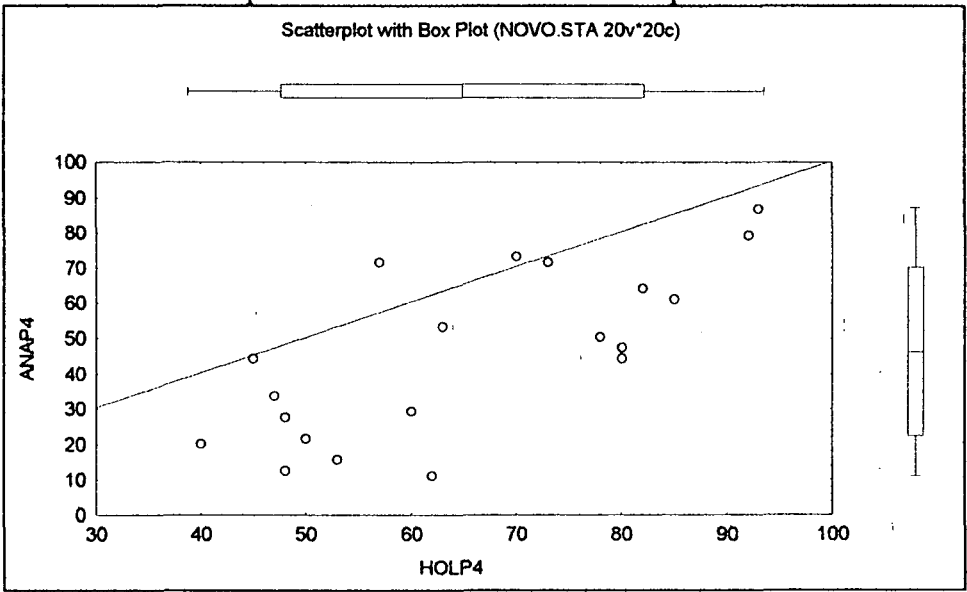


Gráfico 4 - Notas pela Técnica Holística e analítica para o Professor 5

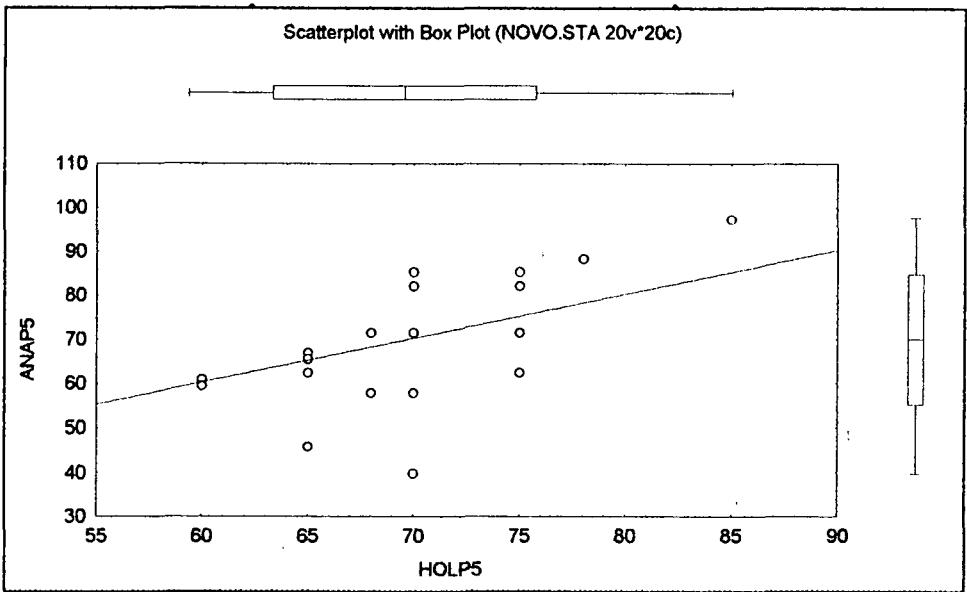


Gráfico 5 - Notas pela Técnica Holística e analítica para o Professor 6

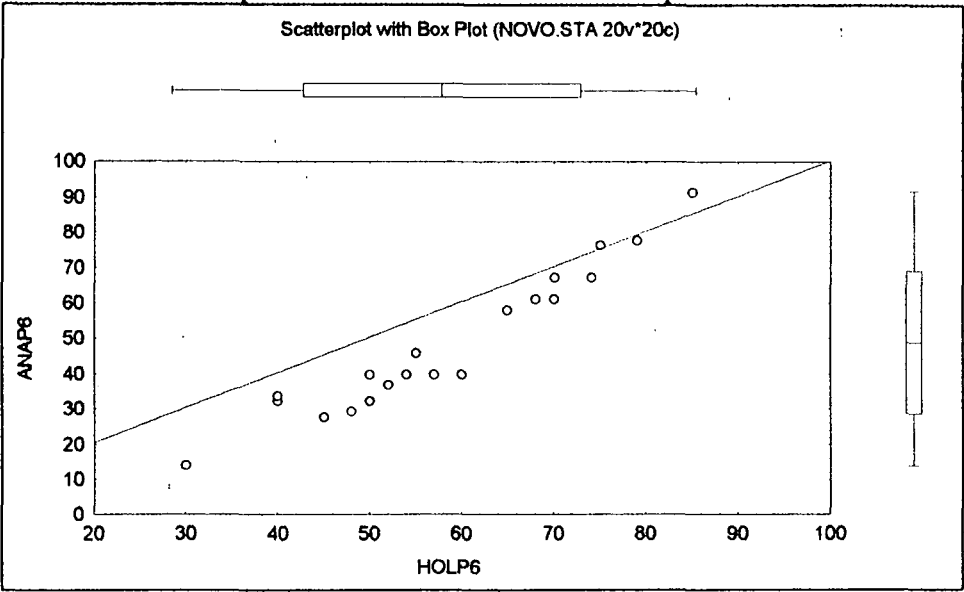


Gráfico 6 - Notas pela Técnica Holística e analítica para o Professor 7

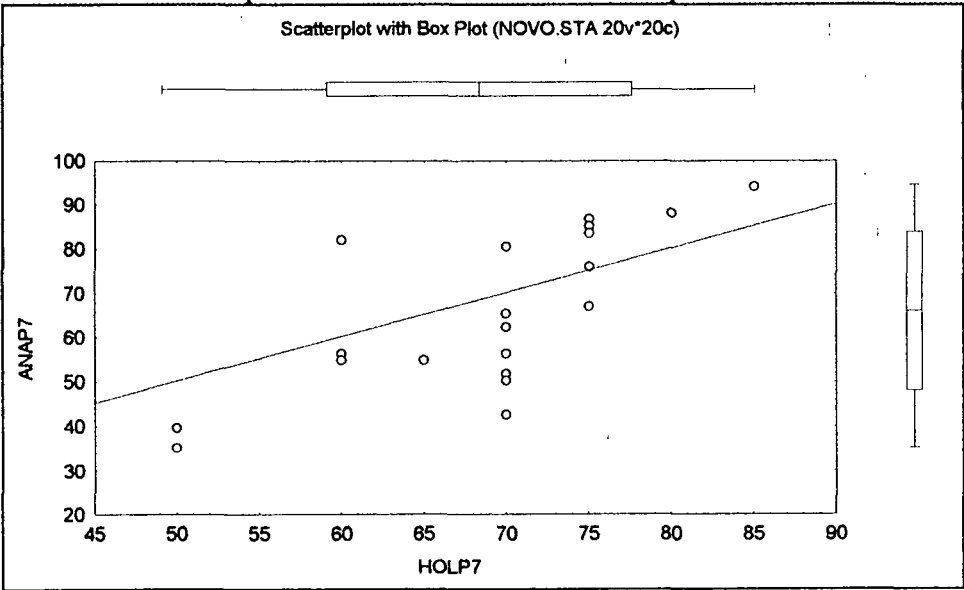


Gráfico 7 - Notas pela Técnica Holística e analítica para o Professor 8

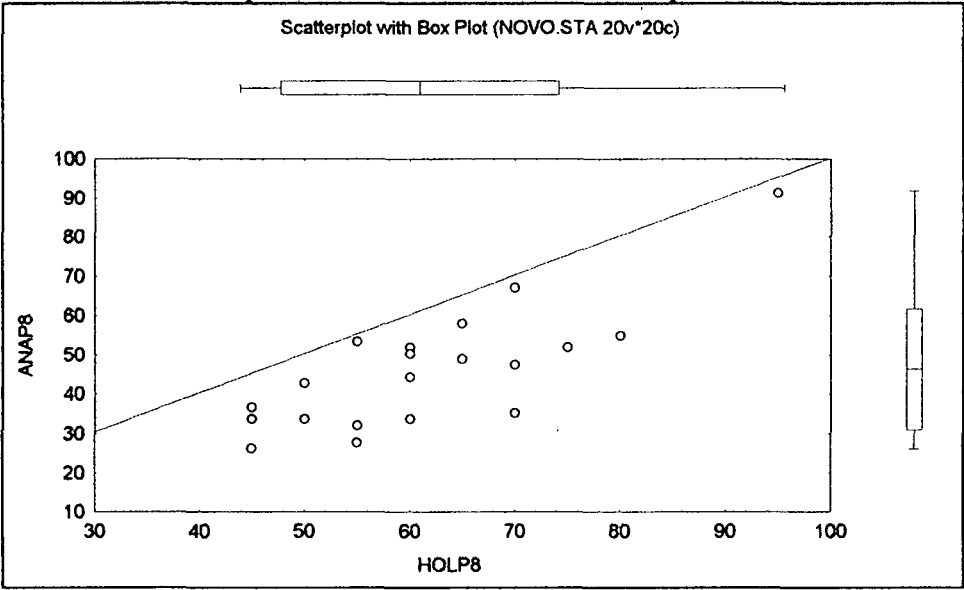


Gráfico 8 - Notas pela Técnica Holística e analítica para o Professor 9

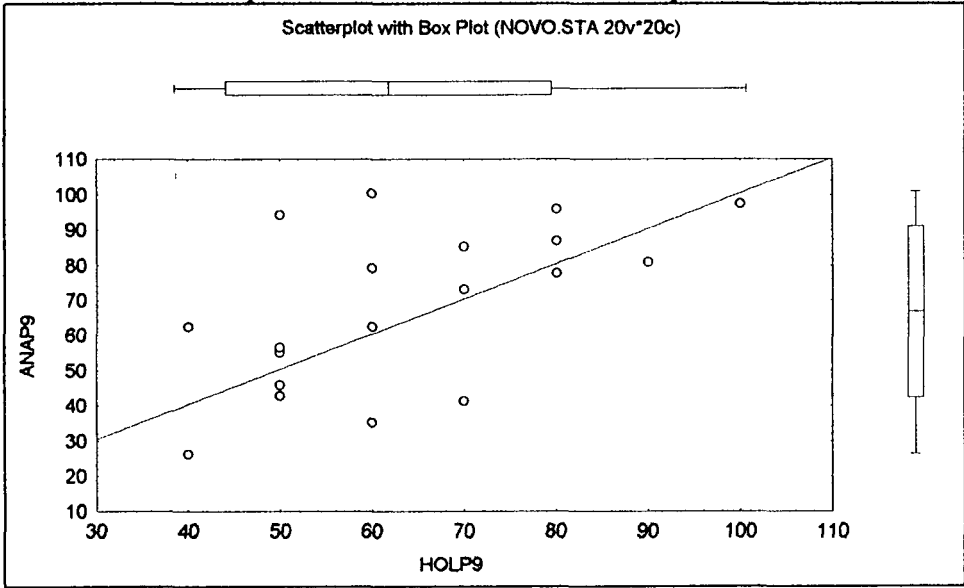
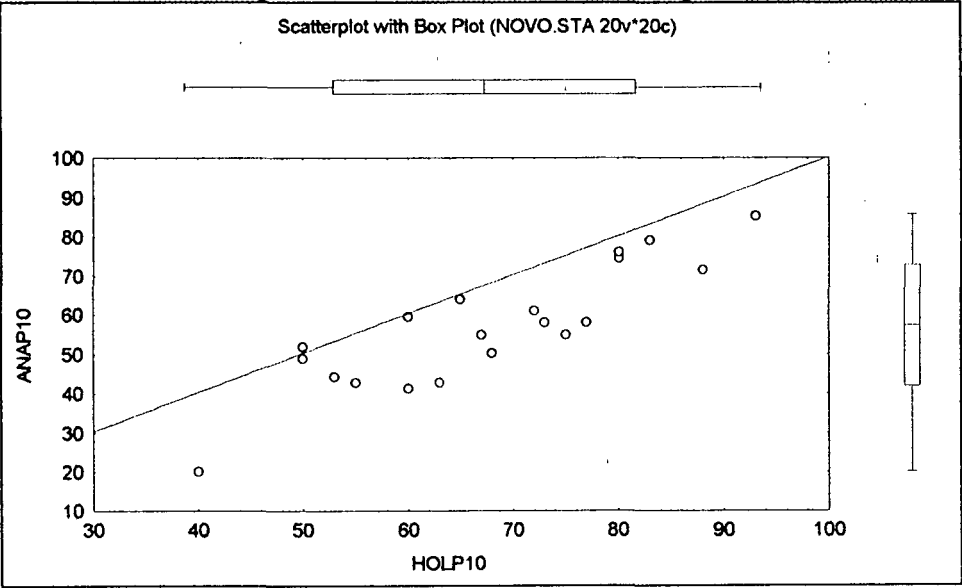


Gráfico 9 - Notas pela Técnica Holística e analítica para o Professor 10



**ANEXO 6 - TABELA DAS MÉDIAS ATRIBUÍDAS PELOS PROFESSORES SEGUNDO A
TÉCNICA HOLÍSTICA**

Tabela das médias atribuídas pelos professores segundo a Técnica Holística

	Prof. 1	Prof. 2	Prof. 3	Prof. 4	Prof. 5	Prof. 6	Prof. 7	Prof. 8	Prof. 9	Prof. 10	Média
R1	30	50	30	62	70	48	70	45	50	72	52,7
R2	40	50	25	78	65	50	50	45	40	68	51,1
R3	40	60	46	80	68	57	70	60	70	60	61,1
R4	70	70	60	70	75	60	75	80	60	83	70,3
R5	60	70	48	80	75	70	70	70	80	67	69
R6	40	45	55	60	75	54	65	60	50	80	58,4
R7	50	70	35	48	70	70	70	65	60	60	59,8
R8	80	75	25	82	70	68	75	65	80	65	68,5
R9	60	80	58	92	70	74	75	70	70	63	71,2
R10	30	50	32	73	75	45	60	55	50	80	55
R11	40	60	37	63	65	55	60	50	70	50	55
R12	40	85	40	57	70	79	75	75	90	75	68,6
R13	40	45	33	47	60	40	70	50	50	50	48,5
R14	45	55	60	50	68	52	70	70	60	53	58,3
R15	10	40	31	48	60	30	50	45	40	40	39,4
R16	50	70	42	40	65	65	75	60	60	55	58,2
R17	30	70	39	53	70	50	70	60	40	77	55,9
R18	40	80	39	85	78	75	80	55	80	88	70
R19	90	90	80	93	85	85	85	95	100	93	89,6
R20	20	50	40	45	65	40	60	55	50	73	49,8
Média	45,25	63,25	42,75	65,3	69,95	58,35	68,75	61,5	62,5	67,6	60,52

**ANEXO 7 - TABELA DAS MÉDIAS ATRIBUÍDAS PELOS PROFESSORES SEGUNDO A
TÉCNICA ANALÍTICA**

Tabela das médias atribuídas pelos professores segundo a Técnica Analítica

	Prof. 1	Prof. 2	Prof. 3	Prof. 4	Prof. 5	Prof. 6	Prof. 7	Prof. 8	Prof. 9	Prof. 10	Média
R1	12,1	47	28,8	10,6	39,4	28,8	51,5	25,8	45,5	60,6	35,01
R2	12,1	51,5	28,8	50	62,1	31,8	34,9	33,3	25,8	50	38,03
R3	10,6	56,1	25,8	43,9	71,2	39,4	65,2	51,5	40,9	59,1	46,37
R4	59,1	69,7	53	72,7	84,9	39,4	66,7	54,6	100	78,8	67,89
R5	31,8	62,1	47	47	81,8	60,6	62,1	47	77,3	54,6	57,13
R6	16,7	60,6	45,5	28,8	62,1	39,4	54,6	50	54,6	74,2	48,65
R7	71,2	68,2	45,5	27,3	57,6	66,7	80,3	48,5	78,8	0,9	54,5
R8	77,3	65,2	42,4	63,6	84,9	60,6	86,4	57,6	86,4	63,6	68,8
R9	63,6	74,2	40,9	78,8	71,2	66,7	84,9	66,7	84,9	42,4	67,43
R10	40,9	74,2	28,8	71,2	71,2	27,3	81,8	31,8	93,9	75,8	59,69
R11	51,5	39,4	31,8	53	66,7	45,5	56,1	42,4	72,7	51,5	51,06
R12	36,4	74,2	62,1	71,2	81,8	77,3	75,8	51,5	80,3	54,6	66,52
R13	31,8	54,6	42,4	33,3	60,6	31,8	56,1	33,3	56,1	48,5	44,85
R14	57,6	69,7	54,6	21,2	57,6	36,4	42,4	34,9	34,9	43,9	45,32
R15	6,1	40,9	40,9	12,1	59,1	13,6	39,4	36,4	25,8	19,7	29,4
R16	57,6	75,8	47	19,7	65,2	57,6	83,3	33,3	62,1	42,4	54,4
R17	37,9	57,6	45,5	15,2	81,8	39,4	50	43,9	62,1	57,6	49,1
R18	33,3	75,8	42,4	60,6	87,9	75,8	87,9	53	95,5	71,2	68,34
R19	92,4	97	86,4	86,4	97	90,9	93,9	90,9	97	84,9	91,68
R20	24,2	59,1	42,4	43,9	45,5	33,3	54,6	27,3	42,4	57,6	43,03
Média	41,2	63,6	44,1	45,5	69,5	48,1	65,4	45,7	65,9	54,6	54,36

ANEXO 8 - MODELO DA ANÁLISE DE VARIÂNCIA

O modelo para este experimento com delineamento Split Plot com três fatores será:

$$Y = \mu + P_i + R_j + M_k + PR_{ij} + PM_{ik} + RM_{jk} + PRM_{ijk} + \varepsilon_{ijk1}$$

onde:

μ = média geral;

P_i = componente associado ao i-ésimo nível do fator P (bloco), σ_p^2 ;

R_j = componente associado ao j-ésimo nível do fator R (efeito aleatório), σ_r^2 ;

M_k = efeito associado ao k-ésimo nível do fator M (efeito fixo), $\phi_m^2 = (\sum M_k^2) / (c - 1)$;

então:

PR_{ij} = tem efeito aleatório, σ_{pr}^2 ,

PM_{ik} = tem efeito aleatório, σ_{pm}^2 ;

RM_{jk} = tem efeito aleatório, σ_{rm}^2 ;

PRM_{ijk} = tem efeito aleatório, σ_{prm}^2 ;

ε_{ijk1} = é o erro experimental com componente igual a σ^2

ANEXO 9 - SOMA DOS QUADRADOS

Soma dos Quadrados

O teste F de Snedecor testará o fator dividindo o seu quadrado médio (QM) pelo correspondente quadrado médio que possibilite analisarmos a influência do termo. Para identificarmos qual o quadrado médio correspondente, é preciso primeiramente, definir alguns conceitos.

Considera-se índice **vivo** aquele presente no termo , índice **ausente** aquele que não existe e índice **morto** aquele presente no termo , porém entre parênteses. Índice morto só aparece no componente do resíduo, onde apenas o último índice é escrito fora dos parênteses.

O número de graus de liberdade (g.l.) para cada termo do modelo é dado pelo produto do número (número de níveis associado a cada termo morto) pelo número de níveis menos um associado a cada termo vivo.

$QM = SQ / g.l.$, ou seja, é a soma do quadrado (SQ) do fator que será testado, dividido pelos graus de liberdade. O cálculo dos quadrados médios é baseado no número de graus de liberdade, como podemos observar a seguir:

Tabela de graus de liberdade e soma do quadrado para os termos do modelo

Termo	Graus de Liberdade	Soma do Quadrado
PPI	$(10-1)=9$	$\sum = (P^2/40) - C$
RRj	$(20-1)=19$	$\sum = (R^2/20) - C$
MMk	$(2-1)=1$	$\sum = (M^2/200) - C$
PPRij	$(10-1)(20-1)=171$	$\sum (PR^2/2) - \sum (P^2/40) - \sum (R^2/20) + C$
PPMik	$(10-1)(2-1)=9$	$\sum (PM^2/20) - \sum (P^2/40) - \sum (M^2/200) + C$
RRMjk	$(20-1)(2-1)=19$	$\sum (RM^2/10) - \sum (R^2/20) - \sum (M^2/200) + C$
PPRMijk	$(10-1)(20-1)(2-1)=171$	$\sum (PRM^2/1) - \sum (R^2/20) - \sum (M^2/200) + \sum (P^2/40) - \sum (PR^2/2) + \sum (PM^2/20) - \sum (RM^2/10) + C$
$\varepsilon_{(ijk)l}$	$10 \cdot 20 \cdot 2(1-1)=0$	$\sum Y_{iikl}^2 - \sum PRM^2$
Total	$10 \cdot 20 \cdot 2 \cdot 1 - 1 = 399$	$\sum Y^2 - C$

Onde C é a correção $C = (\sum Y)^2 / n$, que é o total geral elevado ao quadrado, dividido pelo número de observações.

ANEXO 10 - CÁLCULO DO QUADRADO MÉDIO

Para a estimativa do quadrado médio foi utilizada a Regra Prática sugerida por Montgomery (1991). Vejamos então como ficou a estimativa do quadrado médio do modelo considerado para o nosso experimento:

Quadro de estimativa do quadrado médio segundo a Regra Prática de Montgomery (1991)

Variân- cia	Fonte de Variação	i 10 A	j 20 A	k 2 F	l 1 A	E(QM)
σ_p^2	P_i	1	20	2	1	$\sigma^2 + 2\sigma_{pr}^2 + 40\sigma_p^2$
σ_r^2	R_j	10	1	2	1	$\sigma^2 + 2\sigma_{pr}^2 + 20\sigma_r^2$
σ_{pr}^2	PR_{ij}	1	1	2	1	$\sigma^2 + 2\sigma_{pr}^2$
ϕ_m^2	M_k	10	20	0	1	$\sigma^2 + \sigma_{prm}^2 + 10\sigma_{rm}^2 + 20\sigma_{pm}^2 + 200\phi_m^2$
σ_{pm}^2	PM_{ik}	1	20	0	1	$\sigma^2 + \sigma_{prm}^2 + 20\sigma_{pm}^2$
σ_{rm}^2	RM_{jk}	10	1	0	1	$\sigma^2 + \sigma_{prm}^2 + 10\sigma_{rm}^2$
σ_{prm}^2	PRM_{ijk}	1	1	0	1	$\sigma^2 + \sigma_{prm}^2$
σ^2	ε_{ijkl}	1	1	1	1	σ^2

ANEXO 11 - TESTE F DE SNEDECOR PARA OS TERMOS DO MODELO

Quadro representativo do Teste F de Snedecor para os termos do Modelo

Fonte de Variação	Quadrados Médios	Teste 'F'
P _i	$(\sigma^2 + 2\sigma_{pr}^2 + 40\sigma_p^2) - (\sigma^2 + 2\sigma_{pr}^2)$	QM _A / QM _{AB}
R _j	$(\sigma^2 + 2\sigma_{pr}^2 + 20\sigma_r^2) - (\sigma^2 + 2\sigma_{pr}^2)$	QM _B / QM _{AB}
PR _{ij}	$(\sigma^2 + 2\sigma_{pr}^2) - \sigma^2$	-----
M _k	$(\sigma^2 + \sigma_{prm}^2 + 10\sigma_{rm}^2 + 20\sigma_{pm}^2 + 200\phi_m^2) - [(\sigma^2 + \sigma_{prm}^2 + 20\sigma_{pm}^2) + (\sigma^2 + \sigma_{prm}^2 + 10\sigma_{rm}^2) - (\sigma^2 + \sigma_{prm}^2)]$	QM _C / [QM _{AC} + QM _{BC} - QM _{ABC}]
PM _{ik}	$(\sigma^2 + \sigma_{prm}^2 + 20\sigma_{pm}^2) - (\sigma^2 + \sigma_{prm}^2)$	QM _{AC} / QM _{ABC}
RM _{jk}	$(\sigma^2 + \sigma_{prm}^2 + 10\sigma_{rm}^2) - (\sigma^2 + \sigma_{prm}^2)$	QM _{BC} / QM _{ABC}
PRM _{ijk}	$(\sigma^2 + \sigma_{prm}^2) - \sigma^2$	-----

REFERÊNCIAS BIBLIOGRÁFICAS

- ALDERSON, J. Charles ; HUGHES, Arthur (Eds.) **ELT documents III - Issues in language testing**. London : The British Council, 1980.
- ALDERSON, J. Charles , CLAPHAM, Caroline, WALL, Dianne. **Language test construction and evaluation**. Cambridge : Cambridge University Press, 1995.
- ASPINALL, Patricia; HASHEMI, Louise. **CAE practice tests 1. Teacher's Book**. Cambridge : Cambridge University Press, 1991.
- AVOSANI, Carlos. **As técnicas holística e analítica na avaliação da expressão escrita**. Curitiba, 1981. Tese de Mestrado para obtenção do título de Mestre em Educação. Universidade Federal do Paraná.
- BACHMAN, Lyle F. **Fundamental considerations in language testing**. Oxford : Oxford University Press, 1991.
- BEREITER, C. ; SCARDAMALIA, M. **The psychology of written composition**. New Jersey : The Guildford Press, 1987.
- BLOOMFIELD, Leonard. **Language**. New York : Holt, Rinehart and Winston, 1961 (publicado pela primeira vez em 1933).
- BOWEN, Tim ; MARKS, Jonathan. **Inside teaching**. Oxford : Heinemann, 1994.
- BROOKES, Arthur; GRUNDY, Peter. **Writing for study purposes**. A teacher's guide to developing individual writing skills. Glasgow : Cambridge, 1990.
- BROWN, James Dean. **Understanding Research in Second Language Learning**. New York : Cambridge, 1993.
- BRUMFIT, C.; FINOCCHIARO, M. **The functional-notional approach: from theory to practice**. New York : Oxford University Press, 1983.
- CANDLIN, C. N. Communicative language teaching and the debt to pragmatics. In: RAMEH, C. (Ed.) **Georgetown University Roundtable 1976**. Washington D. C. : Georgetown University Press, 1976.
- CELCE - MURCIA, Marianne. (Ed.). **Teaching English as a second or foreign language**. Rowley : Newbury House, 1979.
- CHOMSKY, Noam. Linguistic theory. In: ALLEN, J. P. B.; VAN BURDEN, A. (Eds.). **Chomsky : selected readings** . London : Oxford, 1966. p. 152-9.
- COE, Norman et al. **Writing skills**. A problem-solving approach. Cambridge : Cambridge, 1994.
- ELBOW, P. **Writing without teachers**. Oxford : Oxford University Press, 1973.

- ELLIS, Rod. **The study of second language acquisition**. Oxford : Oxford, 1994.
- FIRTH, J. R. **Papers in Linguistics: 1934-1951**. London : Oxford University Press, 1957.
- FLOWER, L. S. ; HAYES, JR. A cognitive process theory of writing. **College composition and communication**. 32/4, p. 365-387.
- FRIES, C. C. **Foundations of English teaching**. Ann Harbour : University of Michigan Press, 1945.
- FRIED-BOOTH, Diana L.; HASHEMI, Louise. **PET practice tests 2**. Teacher's Book. Cambridge : Cambridge University Press, 1991.
- GRAVES, D. H. **Writing: teachers and children at work**. New Hampshire : Heinemann, 1983.
- GUMPERZ, J. J.; HYMES, D. (Eds.) **Directions in Sociolinguistics: the Ethnography of communication**. New York : Holt, Rinehart and Winston, 1972.
- HALLIDAY, M. A. K. Language structure and language function. In: LYONS, J. (ed.) **New horizons in Linguistics**. Harmondsworth : Penguin, 1970.
- HAMP-LYONS, Liz. **Newbury House TOEFL preparation Kit**. Preparing for the test of written English. New York : Newbury House, 1989.
- HUGHES, Arthur. **Testing for language teachers**. Cambridge : Cambridge, 1994.
- HYMES, D. On communicative competence. In: PRIDE, J. B. ; HOLMES, J. (Eds.) **Sociolinguistics**. Harmondsworth : Penguin, 1972.
- JACOBS, Holly L. et al. **Teaching ESL composition: principles and techniques**. Rowley : Newbury House, 1983.
- KING, Carol ; INGRAM, Beverly; **From writing to composing**. New York : Cambridge, 1988.
- KROLL, Barbara. (Ed.) **Second language writing**. Research insights for the classroom. New York : Cambridge, 1990.
- LABOV, W. The study of language in its social context. **Studium Generale** , v.23, p. 30-87, 1970.
- LARSEN-FREEMAN, Diane; Long, Michael H. **An introduction to second language acquisition research**. New York : Longman, 1992.
- LEKI, Ilona. Coaching from the margins: issues in written response. In.: KROLL, Barbara. (Ed.) **Second language writing**. Research insights for the classroom. New York : Cambridge, 1990.
- MADSEN, Harold S. **Techniques in testing**. New York : Oxford University Press, 1983.

- MONTGOMERY, D. C. **Design and analysis of experiments**. Canada : John Wiley & Sons, 1991.
- OLLER JR., John. **Language tests in school**. London : Longman, 1986.
- _____. **Issues in language testing research**. Massachusetts : Newbury House, 1983.
- RAIMES, Ann. **Techniques in teaching writing**. Hong Kong : Oxford University Press, 1983.
- RICHARDS, Jack C.; RODGERS, Theodore S. **Approaches and methods in language teaching**. New York : Cambridge, 1986.
- RICHARDS, Jack C.; PLATT, John; PLATT, Heidi. **Dictionary of language teaching & applied linguistics**. Singapore : Longman, 1993.
- RIVERS, Wilga M. **Teaching foreign language skills**. Chicago : University of Chicago Press, 1975.
- SHOHAMY, Elana. **A practical handbook in language testing for the second language teacher**. Tel-Aviv : Tel-Aviv University Press, 1985.
- SOARES, Magda B. ; CAMPOS, Edson N. **Técnicas de redação**. Rio de Janeiro : Ao Livro Técnico, 1979.
- STATIGRAPHICS. Versão 7.1.
- STATISTICA for Windows. Versão 7.0.
- UNIVERSITY of CAMBRIDGE LOCAL EXAMINATION SYNDICATE. **IELTS assessment guide for the writing module**. Cambridge, abril 1995. (Unpublished manual)
- _____. **Cambridge First Certificate : Examination Practice 2**. Teacher's Book. Cambridge : Cambridge University Press, 1986.
- _____. **Cambridge Proficiency Examination Practice 4**. Teacher's Book. Cambridge : Cambridge University Press, 1991.
- WEIR, Cyril J. **Communicative language testing**. Trowbridge : Prentice Hall, 1990.
- _____. **Understanding and developing language tests**. Trowbridge : Prentice Hall, 1993.
- WHITE, Ron ; ARNDT, Valerie. **Process writing**. London : Longman, 1994.
- WIDDOWSON, H. G. The teaching of English as communication. **English language teaching**, v.27, n.1, p. 15-18, 1972.
- ZAMEL, Vivian. Responding to student writing. **TESOL Quarterly**, University of Massachusetts, v. 19, n.1, 79-97, março 1985.
- _____. Writing: the process of discovering meaning. **TESOL Quarterly**, University of Massachusetts, v. 16, n. 2, 195-209, 1982.